

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
15 March 2001 (15.03.2001)

PCT

(10) International Publication Number
WO 01/18042 A2

(51) International Patent Classification⁷: C07K 14/00

(21) International Application Number: PCT/US00/24483

(22) International Filing Date:
6 September 2000 (06.09.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/153,403 10 September 1999 (10.09.1999) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application:
US 60/153,403 (CIP)
Filed on 10 September 1999 (10.09.1999)

(71) Applicant (*for all designated States except US*): INCYTE GENOMICS, INC. [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): TANG, Tom, Y. [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). HILLMAN, Jennifer, L. [US/US]; 230 Monrow Drive, #17, Mountain View, CA 94040 (US). YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). BAUGHN, Mariah, R. [US/US]; 14244 Santiago

Road, San Leandro, CA 94577 (US). AZIMZAI, Yalda [US/US]; 2045 Rock Springs, Hayward, CA 94545 (US). AU-YOUNG, Janice [US/US]; 233 Golden Eagle Lane, Brisbane, CA 94005 (US).

(74) Agents: HAMLET-COX, Diana et al.; Incyte Genomics, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— Without international search report and to be republished upon receipt of that report.

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



WO 01/18042 A2

(54) Title: APOPTOSIS PROTEINS

(57) Abstract: The invention provides human apoptosis proteins (APOP) and polynucleotides which identify and encode APOP. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with expression of APOP.

APOPTOSIS PROTEINS

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of apoptosis proteins and to the
5 use of these sequences in the diagnosis, treatment, and prevention of immune and cell proliferative disorders.

BACKGROUND OF THE INVENTION

Normal development, growth, and homeostasis in multicellular organisms requires a careful
10 balance between the production and destruction of cells in tissues throughout the body. Cell division is a carefully coordinated process with numerous checkpoints and control mechanisms. These mechanisms are designed to regulate DNA replication and to prevent inappropriate or excessive cell proliferation. In contrast, apoptosis is the genetically controlled process by which unneeded or defective cells undergo programmed cell death. Unlike necrotic or injured cells, apoptotic cells are
15 rapidly phagocytosed by neighboring cells or macrophages without leaking their potentially damaging contents into the surrounding tissue or triggering an inflammatory response.

Apoptotic cells undergo distinct morphological changes. Hallmarks of apoptosis include cell shrinkage, nuclear and cytoplasmic condensation, and alterations in plasma membrane topology. Biochemically, apoptotic cells are characterized by increased intracellular calcium concentration, fragmentation of chromosomal DNA into nucleosomal-length units, and expression of novel cell
20 surface components.

Apoptotic events are part of the normal developmental programs of many multicellular organisms. Selective elimination of cells is as important for morphogenesis and tissue remodeling as cell proliferation and differentiation. Apoptosis is also a critical component of the immune response.
25 Immune cells such as cytotoxic T-cells and natural killer cells induce apoptosis in tumor cells or virus-infected cells in vitro. Such activity in vivo may halt the spread of virus or tumor proliferation. In addition, immune cells that fail to distinguish self molecules from foreign molecules must be eliminated to avoid an autoimmune response.

The molecular mechanisms of apoptosis are highly conserved, and many of the key protein
30 regulators and effectors of apoptosis have been identified. Apoptosis generally proceeds in response to a signal which is transduced intracellularly, resulting in altered patterns of gene expression and protein activity. Signaling molecules such as hormones and cytokines are known to regulate apoptosis both positively and negatively. Transcription factors also play an important role in the onset of apoptosis. A number of downstream effector molecules, especially proteases, have been
35 implicated in the degradation of cellular components and the proteolytic activation of other apoptotic

effectors.

The phenomenon of apoptosis involves genes expressed specifically for just such an event. For example, the rat *Rp8* gene is transiently expressed and thought to be involved in programmed cell death (Owens, G.P. and J.J. Cohen (1992) *Cancer Metastasis Rev.* 11:149-156). The human homolog
5 to *Rp8*, *PDCD2*, contains an open reading frame of 1,032 bp, encoding 344 amino acids. *PDCD2* shares 81% identity at the DNA level and 83% identity at the polypeptide level with rat *Rp8*, and is expressed in all tissues examined (Kawakami, T. et al. (1995) *Cytogenet. Cell Genet.* 71:41-43).

The caspase pathway is an example of the use of a signal transduction pathway as an effector arm of the apoptotic program. Caspases are a family of cysteine proteases related to the
10 *Caenorhabditis elegans* CED-3 protein (Alnemri, E.S. et al. (1996) *Cell* 87:171). To date, more than 10 caspases have been identified and partially characterized. Many have been implicated in the induction of apoptosis. Caspases are proteolytically processed from inactive precursors to generate active subunits. They are divided into proximal and distal caspases based on their sites of action and where they reside within the caspase cascade of proteolysis (Icohen, G.M. (1997) *Biochem. J.* 326:1-
15 16). Although little is known regarding regulation of caspase activity during apoptosis, a caspase recruitment domain (CARD), is found within the prodomain of several apical caspases and is conserved in several apoptosis regulatory molecules such as Apaf-2, RAIDD, and cellular inhibitors of apoptosis proteins (IAPs) (Hofmann, K. et al. (1997) *Trends Biochem. Sci.* 22:155-157). The regulatory role in apoptosis of CARD may be to allow proteins such as Apaf-1 to associate with
20 caspase-9 (Li, P. et al. (1997) *Cell* 91:479-489). The regulation of apoptosis in muscle cells is poorly understood. A human cDNA encoding an apoptosis repressor with a CARD (ARC) which is expressed in both skeletal and cardiac muscle has been identified and characterized. ARC functions as an inhibitor of apoptosis and interacts selectively with caspases (Koseki, T. et al. (1998) *Proc. Natl. Acad. Sci. USA* 95:5156-5160).

25 Apoptosis activation also occurs as a result of the induction of gene expression in response to a stimulus. The use of all-trans retinoic acid (ATRA) in the myeloid cell line P39, derived from a patient with myelodysplastic syndrome (MDS), stimulates apoptosis. ATRA-induced apoptosis was mediated through the caspase pathway. The use of retinoic acid as a stimulant of apoptosis in human apoptotic HL-60 cells also detected an apoptosis related protein, APR-1 (Zhu, F. et al. (1999)
30 GenBank Accession No. AF143225.2).

A final step in the apoptotic effector pathway is the fragmentation of nuclear DNA. Nucleolus fragmentation in the yeast *Saccharomyces cerevisiae* is characteristic of cellular aging, and leads to senescence and cell death. The longevity assurance protein 1 (LAG1) of *Saccharomyces cerevisiae* was the first longevity gene to be identified (D'mello, N.P. et al. (1994) *J. Biol. Chem.*
35 269:15451-15459). The *LAG1* gene, when deleted in haploid cells of *Saccharomyces cerevisiae*,

resulted in a pronounced increase in both mean and maximum life span (approximately 50%) of the yeast mother cell. These results indicate that *LAG1* is involved in determination of yeast longevity (D'mello, supra). The gene for a human (*LAG1Hs*) and *Caenorhabditis elegans* (*LAG1Ce-1*) homolog of *LAG1* has been cloned (Jiang, J.C. et al. (1998) Genome Res. 8:1259-1272). Both

5 homologs are able to functionally complement the lethality of a *lag1* double deletion mutant. *LAG1Hs* was able to restore the life span of the double deletion mutant, indicating its function in establishing the longevity phenotype. Expression studies found *LAG1Hs* expressed in only 3 tissues: brain, skeletal muscle, and testis. Although both the human and *C. elegans* proteins have low sequence identity to the yeast *LAG1*, these two proteins share with the yeast protein a short sequence,

10 the *Lag1* motif, and similar transmembrane domain profiles. The expression pattern and ability to complement a yeast *lag1* double deletion mutant suggests that *LAG1Hs* may have a role in human aging (Jaing, supra).

Dysregulation of apoptosis has recently been recognized as a significant factor in the pathogenesis of many human diseases. For example, excessive cell survival caused by decreased

15 apoptosis can contribute to disorders related to cell proliferation and the immune response. Such disorders include cancer, autoimmune diseases, viral infections, and inflammation. In contrast, excessive cell death caused by increased apoptosis can lead to degenerative and immunodeficiency disorders such as AIDS, neurodegenerative diseases, and myelodysplastic syndromes (Thompson, C.B. (1995) Science 267:1456-1462).

20 The discovery of new apoptosis proteins and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of immune and cell proliferative disorders.

SUMMARY OF THE INVENTION

25 The invention features purified polypeptides, apoptosis proteins, referred to collectively as "APOP" and individually as "APOP-1," "APOP-2," "APOP-3," "APOP-4," and "APOP-5." In one aspect, the invention provides an isolated polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid

30 sequence selected from the group consisting of SEQ ID NO:1-5, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-5. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-5.

The invention further provides an isolated polynucleotide encoding a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, c) a
5 biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-5. In one alternative, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-5. In another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:6-10.

10 Additionally, the invention provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, c) a
15 biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-5. In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the invention provides a transgenic organism comprising the recombinant polynucleotide.

20 The invention also provides a method for producing a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID
25 NO:1-5, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-5. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

30 Additionally, the invention provides an isolated antibody which specifically binds to a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, c) a biologically active fragment of an amino acid sequence selected

from the group consisting of SEQ ID NO:1-5, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-5.

The invention further provides an isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of
5 SEQ ID NO:6-10, b) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:6-10, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), and e) an RNA equivalent of a)-d). In one alternative, the polynucleotide comprises at least 60 contiguous nucleotides.

10 Additionally, the invention provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:6-10, b) a naturally occurring polynucleotide sequence having at least 70% sequence
15 identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:6-10, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), and e) an RNA equivalent of a)-d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or
20 fragments thereof, and b) detecting the presence or absence of said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the probe comprises at least 60 contiguous nucleotides.

The invention further provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide comprising a polynucleotide sequence
25 selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:6-10, b) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:6-10, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), and e) an RNA equivalent of a)-d). The method comprises a) amplifying said target polynucleotide or
30 fragment thereof using polymerase chain reaction amplification, and b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

The invention further provides a composition comprising an effective amount of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence

selected from the group consisting of SEQ ID NO:1-5, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, and d) an immunogenic fragment of an amino acid sequence selected
5 from the group consisting of SEQ ID NO:1-5, and a pharmaceutically acceptable excipient. In one embodiment, the composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-5. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional APOP, comprising administering to a patient in need of such treatment the composition.

10 The invention also provides a method for screening a compound for effectiveness as an agonist of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, c) a biologically active fragment of an amino acid sequence
15 selected from the group consisting of SEQ ID NO:1-5, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-5. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting agonist activity in the sample. In one alternative, the invention provides a composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the
20 invention provides a method of treating a disease or condition associated with decreased expression of functional APOP, comprising administering to a patient in need of such treatment the composition.

Additionally, the invention provides a method for screening a compound for effectiveness as an antagonist of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, b) a naturally
25 occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-5. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting
30 antagonist activity in the sample. In one alternative, the invention provides a composition comprising an antagonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with overexpression of functional APOP, comprising administering to a patient in need of such treatment the composition.

35 The invention further provides a method of screening for a compound that specifically binds

to a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, c) a biologically active fragment of an amino acid sequence selected
5 from the group consisting of SEQ ID NO:1-5, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-5. The method comprises a) combining the polypeptide with at least one test compound under suitable conditions, and b) detecting binding of the polypeptide to the test compound, thereby identifying a compound that specifically binds to the polypeptide.

10 The invention further provides a method of screening for a compound that modulates the activity of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, c) a biologically active fragment of an amino
15 acid sequence selected from the group consisting of SEQ ID NO:1-5, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-5. The method comprises a) combining the polypeptide with at least one test compound under conditions permissive for the activity of the polypeptide, b) assessing the activity of the polypeptide in the presence of the test compound, and c) comparing the activity of the polypeptide in the presence of the
20 test compound with the activity of the polypeptide in the absence of the test compound, wherein a change in the activity of the polypeptide in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide.

The invention further provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a
25 sequence selected from the group consisting of SEQ ID NO:6-10, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, and b) detecting altered expression of the target polynucleotide.

The invention further provides a method for assessing toxicity of a test compound, said method comprising a) treating a biological sample containing nucleic acids with the test compound;
30 b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide comprising a polynucleotide sequence selected from the group consisting of i) a polynucleotide sequence selected from the group consisting of SEQ ID NO:6-10, ii) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:6-10, iii) a polynucleotide
35 sequence complementary to i), iv) a polynucleotide sequence complementary to ii), and v) an RNA

equivalent of i)-iv). Hybridization occurs under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence selected from the group consisting of i) a polynucleotide sequence selected from the group consisting of SEQ ID NO:6-10, ii) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:6-10, iii) a polynucleotide sequence complementary to i), iv) a polynucleotide sequence complementary to ii), and v) an RNA equivalent of i)-iv). Alternatively, the target polynucleotide comprises a fragment of a polynucleotide sequence selected from the group consisting of i)-v) above; c) quantifying the amount of hybridization complex; and d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

BRIEF DESCRIPTION OF THE TABLES

Table 1 shows polypeptide and nucleotide sequence identification numbers (SEQ ID NOs), clone identification numbers (clone IDs), cDNA libraries, and cDNA fragments used to assemble full-length sequences encoding APOP.

Table 2 shows features of each polypeptide sequence, including potential motifs, homologous sequences, and methods, algorithms, and searchable databases used for analysis of APOP.

Table 3 shows selected fragments of each nucleic acid sequence; the tissue-specific expression patterns of each nucleic acid sequence as determined by northern analysis; diseases, disorders, or conditions associated with these tissues; and the vector into which each cDNA was cloned.

Table 4 describes the tissues used to construct the cDNA libraries from which cDNA clones encoding APOP were isolated.

Table 5 shows the tools, programs, and algorithms used to analyze the polynucleotides and polypeptides of the invention, along with applicable descriptions, references, and threshold parameters.

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an,"

and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

5 Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines,
10 protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

"APOP" refers to the amino acid sequences of substantially purified APOP obtained from any
15 species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which intensifies or mimics the biological activity of APOP. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of APOP either by directly interacting with
20 APOP or by acting on components of the biological pathway in which APOP participates.

An "allelic variant" is an alternative form of the gene encoding APOP. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. A gene may have none, one, or many allelic variants of its naturally occurring form. Common mutational changes which give rise to
25 allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding APOP include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as APOP or a
30 polypeptide with at least one functional characteristic of APOP. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding APOP, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding APOP. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino

acid residues which produce a silent change and result in a functionally equivalent APOP. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of APOP is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

10 The terms "amino acid" and "amino acid sequence" refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where "amino acid sequence" is recited to refer to a sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

15 "Amplification" relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term "antagonist" refers to a molecule which inhibits or attenuates the biological activity of APOP. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small
20 molecules, or any other compound or composition which modulates the activity of APOP either by directly interacting with APOP or by acting on components of the biological pathway in which APOP participates.

The term "antibody" refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, F(ab')₂, and Fv fragments, which are capable of binding an epitopic determinant.
25 Antibodies that bind APOP polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole
30 limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the
35 protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to

elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition capable of base-pairing with the "sense" (coding) strand of a specific nucleic acid sequence. Antisense compositions may include DNA; RNA; peptide nucleic acid (PNA); oligonucleotides having modified backbone linkages such as
5 phosphorothioates, methylphosphonates, or benzylphosphonates; oligonucleotides having modified sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having modified bases such as 5-methyl cytosine, 2'-deoxyuracil, or 7-deaza-2'-deoxyguanosine. Antisense molecules may be produced by any method including chemical synthesis or transcription. Once introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring
10 nucleic acid sequence produced by the cell to form duplexes which block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand of a reference DNA molecule.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" or "immunogenic"
15 refers to the capability of the natural, recombinant, or synthetic APOP, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

"Complementary" describes the relationship between two single-stranded nucleic acid sequences that anneal by base-pairing. For example, 5'-AGT-3' pairs with its complement,
20 3'-TCA-5'.

A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding APOP or fragments of APOP may be
25 employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been subjected to repeated
30 DNA sequence analysis to resolve uncalled bases, extended using the XL-PCR kit (PE Biosystems, Foster City CA) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from one or more overlapping cDNA, EST, or genomic DNA fragments using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison WI) or Phrap (University of Washington, Seattle WA). Some sequences have been both extended and assembled to

produce the consensus sequence.

“Conservative amino acid substitutions” are those substitutions that are predicted to least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative amino acid substitutions.

	Original Residue	Conservative Substitution
	Ala	Gly, Ser
	Arg	His, Lys
10	Asn	Asp, Gln, His
	Asp	Asn, Glu
	Cys	Ala, Ser
	Gln	Asn, Glu, His
	Glu	Asp, Gln, His
15	Gly	Ala
	His	Asn, Arg, Gln, Glu
	Ile	Leu, Val
	Leu	Ile, Val
	Lys	Arg, Gln, Glu
20	Met	Leu, Ile
	Phe	His, Met, Leu, Trp, Tyr
	Ser	Cys, Thr
	Thr	Ser, Val
	Trp	Phe, Tyr
25	Tyr	His, Phe, Trp
	Val	Ile, Leu, Thr

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

A “deletion” refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term “derivative” refers to a chemically modified polynucleotide or polypeptide. Chemical modifications of a polynucleotide sequence can include, for example, replacement of hydrogen by an alkyl, acyl, hydroxyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

A “detectable label” refers to a reporter molecule or enzyme that is capable of generating a measurable signal and is covalently or noncovalently joined to a polynucleotide or polypeptide.

A "fragment" is a unique portion of APOP or the polynucleotide encoding APOP which is identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment
5 used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 16, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50% of a polypeptide) as shown in a certain
10 defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

A fragment of SEQ ID NO:6-10 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:6-10, for example, as distinct from any other sequence in the
15 genome from which the fragment was obtained. A fragment of SEQ ID NO:6-10 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:6-10 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:6-10 and the region of SEQ ID NO:6-10 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

20 A fragment of SEQ ID NO:1-5 is encoded by a fragment of SEQ ID NO:6-10. A fragment of SEQ ID NO:1-5 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-5. For example, a fragment of SEQ ID NO:1-5 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-5. The precise length of a fragment of SEQ ID NO:1-5 and the region of SEQ ID NO:1-5 to which the fragment corresponds are
25 routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A "full-length" polynucleotide sequence is one containing at least a translation initiation codon (e.g., methionine) followed by an open reading frame and a translation termination codon. A "full-length" polynucleotide sequence encodes a "full-length" polypeptide sequence.

30 "Homology" refers to sequence similarity or, interchangeably, sequence identity, between two or more polynucleotide sequences or two or more polypeptide sequences.

The terms "percent identity" and "% identity," as applied to polynucleotide sequences, refer to the percentage of residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in
35 the sequences being compared in order to optimize alignment between two sequences, and therefore

achieve a more meaningful comparison of the two sequences.

Percent identity between polynucleotide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polynucleotide sequences.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at <http://www.ncbi.nlm.nih.gov/BLAST/>. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at <http://www.ncbi.nlm.nih.gov/gorf/bl2.html>. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62
Reward for match: 1
Penalty for mismatch: -2
Open Gap: 5 and Extension Gap: 2 penalties
Gap x drop-off: 50
Expect: 10
Word Size: 11
Filter: on

Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at

least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

5 Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to
10 the percentage of residue matches between at least two polypeptide sequences aligned using a standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment methods take into account conservative amino acid substitutions. Such conservative substitutions, explained in more detail above, generally preserve the charge and hydrophobicity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide.

15 Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program (described and referenced above). For pairwise alignments of polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default residue weight table. As with
20 polynucleotide alignments, the percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polypeptide sequence pairs.

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.12 (Apr-21-2000) with blastp set at default parameters. Such default parameters may be, for example:

25 *Matrix: BLOSUM62*
Open Gap: 11 and Extension Gap: 1 penalties
Gap x drop-off: 50
Expect: 10
Word Size: 3
30 *Filter: on*

Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150

contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

“Human artificial chromosomes” (HACs) are linear microchromosomes which may contain
5 DNA sequences of about 6 kb to 10 Mb in size, and which contain all of the elements required for chromosome replication, segregation and maintenance.

The term “humanized antibody” refers to an antibody molecule in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

10 “Hybridization” refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of complementarity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the “washing” step(s). The washing step(s) is particularly important in determining the stringency
15 of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive
20 annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 µg/ml sheared, denatured salmon sperm DNA.

Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Such wash temperatures are typically selected to be about 5°C to 20°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic
25 strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating T_m and conditions for nucleic acid hybridization are well known and can be found in Sambrook, J. et al., 1989, Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; specifically see volume 2, chapter 9.

30 High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance,

sheared and denatured salmon sperm DNA at about 100-200 $\mu\text{g/ml}$. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency
5 conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_{0t} or R_{0t} analysis) or formed between one nucleic acid
10 sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" and "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

15 "Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

An "immunogenic fragment" is a polypeptide or oligopeptide fragment of APOP which is
20 capable of eliciting an immune response when introduced into a living organism, for example, a mammal. The term "immunogenic fragment" also includes any polypeptide or oligopeptide fragment of APOP which is useful in any of the antibody production methods disclosed herein or known in the art.

The term "microarray" refers to an arrangement of a plurality of polynucleotides, polypeptides, or other chemical compounds on a substrate.

25 The terms "element" and "array element" refer to a polynucleotide, polypeptide, or other chemical compound having a unique and defined position on a microarray.

The term "modulate" refers to a change in the activity of APOP. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of APOP.

30 The phrases "nucleic acid" and "nucleic acid sequence" refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

"Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a

functional relationship with a second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

5 "Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

10 "Post-translational modification" of an APOP may involve lipidation, glycosylation, phosphorylation, acetylation, racemization, proteolytic cleavage, and other modifications known in the art. These processes may occur synthetically or biochemically. Biochemical modifications will vary by cell type depending on the enzymatic milieu of APOP.

 "Probe" refers to nucleic acid sequences encoding APOP, their complements, or fragments
15 thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. "Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA
20 polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

 Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100,
25 or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the specification, including the tables, figures, and Sequence Listing, may be used.

 Methods for preparing and using probes and primers are described in the references, for example Sambrook, J. et al. (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold
30 Spring Harbor Press, Plainview NY; Ausubel, F.M. et al. (1987) Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY; Innis, M. et al. (1990) PCR Protocols. A Guide to Methods and Applications, Academic Press, San Diego CA. PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge
35 MA).

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

A "recombinant nucleic acid" is a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two or more otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook, supra. The term recombinant includes nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be used to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

A "regulatory element" refers to a nucleic acid sequence usually derived from untranslated

regions of a gene and includes enhancers, promoters, introns, and 5' and 3' untranslated regions (UTRs). Regulatory elements interact with host or viral proteins which control transcription, translation, or RNA stability.

“Reporter molecules” are chemical or biochemical moieties used for labeling a nucleic acid, amino acid, or antibody. Reporter molecules include radionuclides; enzymes; fluorescent, chemiluminescent, or chromogenic agents; substrates; cofactors; inhibitors; magnetic particles; and other moieties known in the art.

An “RNA equivalent,” in reference to a DNA sequence, is composed of the same linear sequence of nucleotides as the reference DNA sequence with the exception that all occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The term “sample” is used in its broadest sense. A sample suspected of containing nucleic acids encoding APOP, or fragments thereof, or APOP itself, may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms “specific binding” and “specifically binding” refer to that interaction between a protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or synthetic binding composition. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope “A,” the presence of a polypeptide comprising the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term “substantially purified” refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least 60% free, preferably at least 75% free, and most preferably at least 90% free from other components with which they are naturally associated.

A “substitution” refers to the replacement of one or more amino acid residues or nucleotides by different amino acid residues or nucleotides, respectively.

“Substrate” refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

A “transcript image” refers to the collective pattern of gene expression by a particular cell type or tissue under given conditions at a given time.

“Transformation” describes a process by which exogenous DNA is introduced into a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, bacteriophage or viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term “transformed” cells includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

10 A “transgenic organism,” as used herein, is any organism, including but not limited to animals and plants, in which one or more of the cells of the organism contains heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or in vitro fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The transgenic organisms contemplated in accordance with the present invention include bacteria, cyanobacteria, fungi, plants, and animals. The isolated DNA of the present invention can be introduced into the host by methods known in the art, for example infection, transfection, transformation or transconjugation. Techniques for transferring the DNA of the present invention into such organisms are widely known and provided in references such as Sambrook, J. et al. (1989), supra.

A “variant” of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the “BLAST 2 Sequences” tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95% or at least 98% or greater sequence identity over a certain defined length. A variant may be described as, for example, an “allelic” (as defined above), “splice,” “species,” or “polymorphic” variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternative splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given

species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 98% or greater sequence identity over a certain defined length of one of the polypeptides.

10 THE INVENTION

The invention is based on the discovery of new human apoptosis proteins (APOP), the polynucleotides encoding APOP, and the use of these compositions for the diagnosis, treatment, or prevention of immune and cell proliferative disorders.

Table 1 lists the Incyte clones used to assemble full length nucleotide sequences encoding APOP. Columns 1 and 2 show the sequence identification numbers (SEQ ID NOs) of the polypeptide and nucleotide sequences, respectively. Column 3 shows the clone IDs of the Incyte clones in which nucleic acids encoding each APOP were identified, and column 4 shows the cDNA libraries from which these clones were isolated. Column 5 shows Incyte clones and their corresponding cDNA libraries. Clones for which cDNA libraries are not indicated were derived from pooled cDNA libraries. The Incyte clones in column 5 were used to assemble the consensus nucleotide sequence of each APOP and are useful as fragments in hybridization technologies.

The columns of Table 2 show various properties of each of the polypeptides of the invention: column 1 references the SEQ ID NO; column 2 shows the number of amino acid residues in each polypeptide; column 3 shows potential phosphorylation sites; column 4 shows potential glycosylation sites; column 5 shows the amino acid residues comprising signature sequences and motifs; column 6 shows homologous sequences as identified by BLAST analysis; and column 7 shows analytical methods and in some cases, searchable databases to which the analytical methods were applied. The methods of column 7 were used to characterize each polypeptide through sequence homology and protein motifs.

The columns of Table 3 show the tissue-specificity and diseases, disorders, or conditions associated with nucleotide sequences encoding APOP. The first column of Table 3 lists the nucleotide SEQ ID NOs. Column 2 lists fragments of the nucleotide sequences of column 1. These fragments are useful, for example, in hybridization or amplification technologies to identify SEQ ID NO:6-10 and to distinguish between SEQ ID NO:6-10 and related polynucleotide sequences. The polypeptides encoded by these fragments are useful, for example, as immunogenic peptides. Column 3 lists tissue categories which express APOP as a fraction of total tissues expressing APOP. Column 4 lists

diseases, disorders, or conditions associated with those tissues expressing APOP as a fraction of total tissues expressing APOP. Column 5 lists the vectors used to subclone each cDNA library.

The columns of Table 4 show descriptions of the tissues used to construct the cDNA libraries from which cDNA clones encoding APOP were isolated. Column 1 references the nucleotide SEQ ID NOs, column 2 shows the cDNA libraries from which these clones were isolated, and column 3 shows the tissue origins and other descriptive information relevant to the cDNA libraries in column 2.

SEQ ID NO:6 maps to chromosome 7 within the interval from 84.90 to 90.30 centiMorgans. SEQ ID NO:8 maps to chromosome X within the interval from 68.20 to 90.80 centiMorgans.

The invention also encompasses APOP variants. A preferred APOP variant is one which has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid sequence identity to the APOP amino acid sequence, and which contains at least one functional or structural characteristic of APOP.

The invention also encompasses polynucleotides which encode APOP. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:6-10, which encodes APOP. The polynucleotide sequences of SEQ ID NO:6-10, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The invention also encompasses a variant of a polynucleotide sequence encoding APOP. In particular, such a variant polynucleotide sequence will have at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding APOP. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:6-10 which has at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:6-10. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of APOP.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding APOP, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring APOP, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode APOP and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring APOP under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding APOP or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring
5 codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding APOP and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced
10 from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode APOP and APOP derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into
15 a sequence encoding APOP or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:6-10 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) *Methods Enzymol.* 152:399-407; Kimmel, A.R. (1987) *Methods Enzymol.*
20 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (PE Biosystems,
25 Foster City CA), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (PE
30 Biosystems). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (PE Biosystems), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular

Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding APOP may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, 5 restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Applic. 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids 10 Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences 15 are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060.) Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences, 20 Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include 25 sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary 30 sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, PE Biosystems), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled.

Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode APOP may be cloned in recombinant DNA molecules that direct expression of APOP, or
5 fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express APOP.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter APOP-encoding sequences for a variety of purposes including, but not
10 limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

15 The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent Number 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Cramer, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve the biological properties of APOP, such as its biological or enzymatic activity or its ability
20 to bind to other molecules or compounds. DNA shuffling is a process by which a library of gene variants is produced using PCR-mediated recombination of gene fragments. The library is then subjected to selection or screening procedures that identify those gene variants with the desired properties. These preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial"
25 breeding and rapid molecular evolution. For example, fragments of a single gene containing random point mutations may be recombined, screened, and then reshuffled until the desired properties are optimized. Alternatively, fragments of a given gene may be recombined with fragments of homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable
30 manner.

In another embodiment, sequences encoding APOP may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucleic Acids Symp. Ser. 7:215-223; Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232.) Alternatively, APOP itself or a fragment thereof may be synthesized using chemical methods. For example, peptide
35 synthesis can be performed using various solution-phase or solid-phase techniques. (See, e.g.,

Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY, pp. 55-60; and Roberge, J.Y. et al. (1995) *Science* 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (PE Biosystems). Additionally, the amino acid sequence of APOP, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide or a polypeptide having a sequence of a naturally occurring polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g., Chiez, R.M. and F.Z. Regnier (1990) *Methods Enzymol.* 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, supra, pp. 28-53.)

In order to express a biologically active APOP, the nucleotide sequences encoding APOP or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding APOP. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding APOP. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding APOP and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) *Results Probl. Cell Differ.* 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding APOP and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding APOP. These include, but are not limited to, microorganisms such as bacteria transformed

with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or
5 animal cell systems. (See, e.g., Sambrook, supra; Ausubel, supra; Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509; Bitter, G.A. et al. (1987) Methods Enzymol. 153:516-544; Scorer, C.A. et al. (1994) Bio/Technology 12:181-184; Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945; Takamatsu, N. (1987) EMBO J. 6:307-311; Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al.
10 (1984) Science 224:838-843; Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105; The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196; Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659; and Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.) Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for
15 delivery of nucleotide sequences to the targeted organ, tissue, or cell population. (See, e.g., Di Nicola, M. et al. (1998) Cancer Gen. Ther. 5(6):350-356; Yu, M. et al. (1993) Proc. Natl. Acad. Sci. USA 90(13):6340-6344; Buller, R.M. et al. (1985) Nature 317(6040):813-815; McGregor, D.P. et al. (1994) Mol. Immunol. 31(3):219-226; and Verma, I.M. and N. Somia (1997) Nature 389:239-242.) The invention is not limited by the host cell employed.

20 In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding APOP. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding APOP can be achieved using a multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding APOP into the vector's multiple cloning site
25 disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of APOP are needed, e.g. for the production of antibodies,
30 vectors which direct high level expression of APOP may be used. For example, vectors containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of APOP. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such
35 vectors direct either the secretion or intracellular retention of expressed proteins and enable integration

of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra; Bitter, supra; and Scorer, supra.)

Plant systems may also be used for expression of APOP. Transcription of sequences encoding APOP may be driven viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in
5 combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, supra; Broglie, supra; and Winter, supra.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g.,
10 The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding APOP may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain
15 infective virus which expresses APOP in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of
20 DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of APOP in cell lines is preferred. For example, sequences encoding APOP can be transformed into cell
25 lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the
30 introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk*⁻ and *apr*⁻ cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232;

- Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), β glucuronidase and its substrate β -glucuronide, or luciferase and its substrate luciferin may be used.
- 10 These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding APOP is inserted within a marker gene sequence, transformed cells containing sequences encoding APOP can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding APOP under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

20 In general, host cells that contain the nucleic acid sequence encoding APOP and that express APOP may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

25 Immunological methods for detecting and measuring the expression of APOP using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on APOP is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods. a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ.)

30

A wide variety of labels and conjugation techniques are known by those skilled in the art and

may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding APOP include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding APOP, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding APOP may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode APOP may be designed to contain signal sequences which direct secretion of APOP through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding APOP may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric APOP protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of APOP activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and

hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the APOP encoding sequence and the heterologous protein sequence, so that APOP may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, supra, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

10 In a further embodiment of the invention, synthesis of radiolabeled APOP may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, ³⁵S-methionine.

15 APOP of the present invention or fragments thereof may be used to screen for compounds that specifically bind to APOP. At least one and up to a plurality of test compounds may be screened for specific binding to APOP. Examples of test compounds include antibodies, oligonucleotides, proteins (e.g., receptors), or small molecules.

In one embodiment, the compound thus identified is closely related to the natural ligand of APOP, e.g., a ligand or fragment thereof, a natural substrate, a structural or functional mimetic, or a natural binding partner. (See, e.g., Coligan, J.E. et al. (1991) Current Protocols in Immunology 1(2): Chapter 5.) Similarly, the compound can be closely related to the natural receptor to which APOP binds, or to at least a fragment of the receptor, e.g., the ligand binding site. In either case, the compound can be rationally designed using known techniques. In one embodiment, screening for these compounds involves producing appropriate cells which express APOP, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Drosophila, or E. coli. Cells expressing APOP or cell membrane fractions which contain APOP are then contacted with a test compound and binding, stimulation, or inhibition of activity of either APOP or the compound is analyzed.

30 An assay may simply test binding of a test compound to the polypeptide, wherein binding is detected by a fluorophore, radioisotope, enzyme conjugate, or other detectable label. For example, the assay may comprise the steps of combining at least one test compound with APOP, either in solution or affixed to a solid support, and detecting the binding of APOP to the compound. Alternatively, the assay may detect or measure binding of a test compound in the presence of a labeled competitor. Additionally, the assay may be carried out using cell-free preparations, chemical

35

libraries, or natural product mixtures, and the test compound(s) may be free in solution or affixed to a solid support.

APOP of the present invention or fragments thereof may be used to screen for compounds that modulate the activity of APOP. Such compounds may include agonists, antagonists, or partial or
5 inverse agonists. In one embodiment, an assay is performed under conditions permissive for APOP activity, wherein APOP is combined with at least one test compound, and the activity of APOP in the presence of a test compound is compared with the activity of APOP in the absence of the test compound. A change in the activity of APOP in the presence of the test compound is indicative of a compound that modulates the activity of APOP. Alternatively, a test compound is combined with an
10 in vitro or cell-free system comprising APOP under conditions suitable for APOP activity, and the assay is performed. In either of these assays, a test compound which modulates the activity of APOP may do so indirectly and need not come in direct contact with the test compound. At least one and up to a plurality of test compounds may be screened.

In another embodiment, polynucleotides encoding APOP or their mammalian homologs may
15 be "knocked out" in an animal model system using homologous recombination in embryonic stem (ES) cells. Such techniques are well known in the art and are useful for the generation of animal models of human disease. (See, e.g., U.S. Patent No. 5,175,383 and U.S. Patent No. 5,767,337.) For example, mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and grown in culture. The ES cells are transformed with a vector containing the gene of
20 interest disrupted by a marker gene, e.g., the neomycin phosphotransferase gene (neo; Capecchi, M.R. (1989) Science 244:1288-1292). The vector integrates into the corresponding region of the host genome by homologous recombination. Alternatively, homologous recombination takes place using the Cre-loxP system to knockout a gene of interest in a tissue- or developmental stage-specific manner (Marth, J.D. (1996) Clin. Invest. 97:1999-2002; Wagner, K.U. et al. (1997) Nucleic Acids
25 Res. 25:4323-4330). Transformed ES cells are identified and microinjected into mouse cell blastocysts such as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred to pseudopregnant dams, and the resulting chimeric progeny are genotyped and bred to produce heterozygous or homozygous strains. Transgenic animals thus generated may be tested with potential therapeutic or toxic agents.

30 Polynucleotides encoding APOP may also be manipulated in vitro in ES cells derived from human blastocysts. Human ES cells have the potential to differentiate into at least eight separate cell lineages including endoderm, mesoderm, and ectodermal cell types. These cell lineages differentiate into, for example, neural cells, hematopoietic lineages, and cardiomyocytes (Thomson, J.A. et al. (1998) Science 282:1145-1147).

35 Polynucleotides encoding APOP can also be used to create "knockin" humanized animals

(pigs) or transgenic animals (mice or rats) to model human disease. With knockin technology, a region of a polynucleotide encoding APOP is injected into animal ES cells, and the injected sequence integrates into the animal cell genome. Transformed cells are injected into blastulae, and the blastulae are implanted as described above. Transgenic progeny or inbred lines are studied and

5 treated with potential pharmaceutical agents to obtain information on treatment of a human disease. Alternatively, a mammal inbred to overexpress APOP, e.g., by secreting APOP in its milk, may also serve as a convenient source of that protein (Janne, J. et al. (1998) *Biotechnol. Annu. Rev.* 4:55-74).

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists

10 between regions of APOP and apoptosis proteins. In addition, the expression of APOP is closely associated with cancer, cell proliferation, and the immune response. Therefore, APOP appears to play a role in immune and cell proliferative disorders. In the treatment of disorders associated with increased APOP expression or activity, it is desirable to decrease the expression or activity of APOP. In the treatment of disorders associated with decreased APOP expression or activity, it is desirable to

15 increase the expression or activity of APOP.

Therefore, in one embodiment, APOP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of APOP. Examples of such disorders include, but are not limited to, an immune disorder such as inflammation, actinic keratosis, acquired immunodeficiency syndrome (AIDS), Addison's

20 disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, arteriosclerosis, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, bronchitis, bursitis, cholecystitis, cirrhosis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's

25 thyroiditis, paroxysmal nocturnal hemoglobinuria, hepatitis, hypereosinophilia, irritable bowel syndrome, episodic lymphopenia with lymphocytotoxins, mixed connective tissue disease (MCTD), multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, myelofibrosis, osteoarthritis, osteoporosis, pancreatitis, polycythemia vera, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic

30 lupus erythematosus, systemic sclerosis, primary thrombocythemia, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, trauma, and hematopoietic cancer including lymphoma, leukemia, and myeloma; and a cell proliferative disorder such as actinic keratosis, paroxysmal nocturnal hemoglobinuria, and cancers including adenocarcinoma, melanoma, sarcoma, teratocarcinoma, and, in

35 particular, a cancer of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall

bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus.

In another embodiment, a vector capable of expressing APOP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased
5 expression or activity of APOP including, but not limited to, those described above.

In a further embodiment, a composition comprising a substantially purified APOP in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of APOP including, but not limited to, those provided above.

10 In still another embodiment, an agonist which modulates the activity of APOP may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of APOP including, but not limited to, those listed above.

In a further embodiment, an antagonist of APOP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of APOP. Examples of such
15 disorders include, but are not limited to, those immune and cell proliferative disorders described above. In one aspect, an antibody which specifically binds APOP may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express APOP.

In an additional embodiment, a vector expressing the complement of the polynucleotide
20 encoding APOP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of APOP including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by
25 one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of APOP may be produced using methods which are generally known in the art.
30 In particular, purified APOP may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind APOP. Antibodies to APOP may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are

generally preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with APOP or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to APOP have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein. Short stretches of APOP amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to APOP may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030; and Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce APOP-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for APOP may also be generated. For

example, such fragments include, but are not limited to, $F(ab')_2$ fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the $F(ab')_2$ fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. 5 (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between APOP and its specific 10 antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering APOP epitopes is generally used, but a competitive binding assay may also be employed (Pound, supra).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for APOP. Affinity is expressed as an association 15 constant, K_a , which is defined as the molar concentration of APOP-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple APOP epitopes, represents the average affinity, or avidity, of the antibodies for APOP. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular APOP epitope, 20 represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} L/mole are preferred for use in immunoassays in which the APOP-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7 L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of APOP, preferably in active form, from the antibody (Catty, D. (1988) 25 Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC; Liddell, J.E. and A. Cryer (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg 30 specific antibody/ml, is generally employed in procedures requiring precipitation of APOP-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al., supra.)

In another embodiment of the invention, the polynucleotides encoding APOP, or any fragment

or complement thereof, may be used for therapeutic purposes. In one aspect, modifications of gene expression can be achieved by designing complementary sequences or antisense molecules (DNA, RNA, PNA, or modified oligonucleotides) to the coding or regulatory regions of the gene encoding APOP. Such technology is well known in the art, and antisense oligonucleotides or larger fragments can be
5 designed from various locations along the coding or control regions of sequences encoding APOP. (See, e.g., Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press Inc., Totawa NJ.)

In therapeutic use, any gene delivery system suitable for introduction of the antisense sequences into appropriate target cells can be used. Antisense sequences can be delivered intracellularly in the form of an expression plasmid which, upon transcription, produces a sequence
10 complementary to at least a portion of the cellular sequence encoding the target protein. (See, e.g., Slater, J.E. et al. (1998) *J. Allergy Clin. Immunol.* 102(3):469-475; and Scanlon, K.J. et al. (1995) 9(13):1288-1296.) Antisense sequences can also be introduced intracellularly through the use of viral vectors, such as retrovirus and adeno-associated virus vectors. (See, e.g., Miller, A.D. (1990) *Blood* 76:271; Ausubel, supra; Uckert, W. and W. Walther (1994) *Pharmacol. Ther.* 63(3):323-347.) Other
15 gene delivery mechanisms include liposome-derived systems, artificial viral envelopes, and other systems known in the art. (See, e.g., Rossi, J.J. (1995) *Br. Med. Bull.* 51(1):217-225; Boado, R.J. et al. (1998) *J. Pharm. Sci.* 87(11):1308-1315; and Morris, M.C. et al. (1997) *Nucleic Acids Res.* 25(14):2730-2736.)

In another embodiment of the invention, polynucleotides encoding APOP may be used for
20 somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency (e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-linked inheritance (Cavazzana-Calvo, M. et al. (2000) *Science* 288:669-672), severe combined immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency (Blaese, R.M. et al. (1995) *Science* 270:475-480; Bordignon, C. et al. (1995) *Science* 270:470-475),
25 cystic fibrosis (Zabner, J. et al. (1993) *Cell* 75:207-216; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:643-666; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:667-703), thalassemias, familial hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) *Science* 270:404-410; Verma, I.M. and N. Somia (1997) *Nature* 389:239-242)), (ii) express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated
30 cell proliferation), or (iii) express a protein which affords protection against intracellular parasites (e.g., against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D. (1988) *Nature* 335:395-396; Poeschla, E. et al. (1996) *Proc. Natl. Acad. Sci. USA.* 93:11395-11399), hepatitis B or C virus (HBV, HCV); fungal parasites, such as Candida albicans and Paracoccidioides brasiliensis; and protozoan parasites such as Plasmodium falciparum and Trypanosoma cruzi). In the
35 case where a genetic deficiency in APOP expression or regulation causes disease, the expression of

APOP from an appropriate population of transduced cells may alleviate the clinical manifestations caused by the genetic deficiency.

In a further embodiment of the invention, diseases or disorders caused by deficiencies in APOP are treated by constructing mammalian expression vectors encoding APOP and introducing these
5 vectors by mechanical means into APOP-deficient cells. Mechanical transfer technologies for use with cells in vivo or ex vitro include (i) direct DNA microinjection into individual cells, (ii) ballistic gold particle delivery, (iii) liposome-mediated transfection, (iv) receptor-mediated gene transfer, and (v) the use of DNA transposons (Morgan, R.A. and W.F. Anderson (1993) *Annu. Rev. Biochem.* 62:191-217; Ivics, Z. (1997) *Cell* 91:501-510; Boulay, J-L. and H. R  cipon (1998) *Curr. Opin. Biotechnol.* 9:445-
10 450).

Expression vectors that may be effective for the expression of APOP include, but are not limited to, the PCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX vectors (Invitrogen, Carlsbad CA), PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF, PTET-ON, PTRE2, PTRE2-LUC, PTK-HYG (Clontech, Palo Alto CA). APOP may be expressed
15 using (i) a constitutively active promoter, (e.g., from cytomegalovirus (CMV), Rous sarcoma virus (RSV), SV40 virus, thymidine kinase (TK), or β -actin genes), (ii) an inducible promoter (e.g., the tetracycline-regulated promoter (Gossen, M. and H. Bujard (1992) *Proc. Natl. Acad. Sci. USA* 89:5547-5551; Gossen, M. et al. (1995) *Science* 268:1766-1769; Rossi, F.M.V. and H.M. Blau (1998) *Curr. Opin. Biotechnol.* 9:451-456), commercially available in the T-REX plasmid (Invitrogen)); the
20 ecdysone-inducible promoter (available in the plasmids PVGRXR and PIND; Invitrogen); the FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V. and H.M. Blau, *supra*)), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding APOP from a normal individual.

Commercially available liposome transformation kits (e.g., the PERFECT LIPID
25 TRANSFECTION KIT, available from Invitrogen) allow one with ordinary skill in the art to deliver polynucleotides to target cells in culture and require minimal effort to optimize experimental parameters. In the alternative, transformation is performed using the calcium phosphate method (Graham, F.L. and A.J. Eb (1973) *Virology* 52:456-467), or by electroporation (Neumann, E. et al. (1982) *EMBO J.* 1:841-845). The introduction of DNA to primary cells requires modification of these
30 standardized mammalian transfection protocols.

In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to APOP expression are treated by constructing a retrovirus vector consisting of (i) the polynucleotide encoding APOP under the control of an independent promoter or the retrovirus long terminal repeat (LTR) promoter, (ii) appropriate RNA packaging signals, and (iii) a Rev-responsive

element (RRE) along with additional retrovirus *cis*-acting RNA sequences and coding sequences required for efficient vector propagation. Retrovirus vectors (e.g., PFB and PFBNEO) are commercially available (Stratagene) and are based on published data (Riviere, I. et al. (1995) Proc. Natl. Acad. Sci. USA 92:6733-6737), incorporated by reference herein. The vector is propagated in an appropriate vector producing cell line (VPCL) that expresses an envelope gene with a tropism for receptors on the target cells or a promiscuous envelope protein such as VSVg (Armentano, D. et al. (1987) J. Virol. 61:1647-1650; Bender, M.A. et al. (1987) J. Virol. 61:1639-1646; Adam, M.A. and A.D. Miller (1988) J. Virol. 62:3802-3806; Dull, T. et al. (1998) J. Virol. 72:8463-8471; Zufferey, R. et al. (1998) J. Virol. 72:9873-9880). U.S. Patent Number 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant") discloses a method for obtaining retrovirus packaging cell lines and is hereby incorporated by reference. Propagation of retrovirus vectors, transduction of a population of cells (e.g., CD4⁺ T-cells), and the return of transduced cells to a patient are procedures well known to persons skilled in the art of gene therapy and have been well documented (Ranga, U. et al. (1997) J. Virol. 71:7020-7029; Bauer, G. et al. (1997) Blood 89:2259-2267; Bonyhadi, M.L. (1997) J. Virol. 71:4707-4716; Ranga, U. et al. (1998) Proc. Natl. Acad. Sci. USA 95:1201-1206; Su, L. (1997) Blood 89:2283-2290).

In the alternative, an adenovirus-based gene therapy delivery system is used to deliver polynucleotides encoding APOP to cells which have one or more genetic abnormalities with respect to the expression of APOP. The construction and packaging of adenovirus-based vectors are well known to those with ordinary skill in the art. Replication defective adenovirus vectors have proven to be versatile for importing genes encoding immunoregulatory proteins into intact islets in the pancreas (Csete, M.E. et al. (1995) Transplantation 27:263-268). Potentially useful adenoviral vectors are described in U.S. Patent Number 5,707,618 to Armentano ("Adenovirus vectors for gene therapy"), hereby incorporated by reference. For adenoviral vectors, see also Antinozzi, P.A. et al. (1999) Annu. Rev. Nutr. 19:511-544; and Verma, I.M. and N. Somia (1997) Nature 18:389:239-242, both incorporated by reference herein.

In another alternative, a herpes-based, gene therapy delivery system is used to deliver polynucleotides encoding APOP to target cells which have one or more genetic abnormalities with respect to the expression of APOP. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing APOP to cells of the central nervous system, for which HSV has a tropism. The construction and packaging of herpes-based vectors are well known to those with ordinary skill in the art. A replication-competent herpes simplex virus (HSV) type 1-based vector has been used to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) Exp. Eye Res. 169:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S.

Patent Number 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent Number 5,804,413 teaches the use of recombinant HSV d92 which consists of a genome containing at least one exogenous gene to be transferred to a cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this

5 patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 and ICP22. For HSV vectors, see also Goins, W.F. et al. (1999) J. Virol. 73:519-532 and Xu, H. et al. (1994) Dev. Biol. 163:152-161, hereby incorporated by reference. The manipulation of cloned herpesvirus sequences, the generation of recombinant virus following the transfection of multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation of

10 herpesvirus, and the infection of cells with herpesvirus are techniques well known to those of ordinary skill in the art.

In another alternative, an alphavirus (positive, single-stranded RNA virus) vector is used to deliver polynucleotides encoding APOP to target cells. The biology of the prototypic alphavirus, Semliki Forest Virus (SFV), has been studied extensively and gene transfer vectors have been based on

15 the SFV genome (Garoff, H. and K.-J. Li (1998) Curr. Opin. Biotechnol. 9:464-469). During alphavirus RNA replication, a subgenomic RNA is generated that normally encodes the viral capsid proteins. This subgenomic RNA replicates to higher levels than the full-length genomic RNA, resulting in the overproduction of capsid proteins relative to the viral proteins with enzymatic activity (e.g., protease and polymerase). Similarly, inserting the coding sequence for APOP into the alphavirus

20 genome in place of the capsid-coding region results in the production of a large number of APOP-coding RNAs and the synthesis of high levels of APOP in vector transduced cells. While alphavirus infection is typically associated with cell lysis within a few days, the ability to establish a persistent infection in hamster normal kidney cells (BHK-21) with a variant of Sindbis virus (SIN) indicates that the lytic replication of alphaviruses can be altered to suit the needs of the gene therapy application

25 (Dryga, S.A. et al. (1997) Virology 228:74-83). The wide host range of alphaviruses will allow the introduction of APOP into a variety of cell types. The specific transduction of a subset of cells in a population may require the sorting of cells prior to transduction. The methods of manipulating infectious cDNA clones of alphaviruses, performing alphavirus cDNA and RNA transfections, and performing alphavirus infections, are well known to those with ordinary skill in the art.

30 Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may also be employed to inhibit gene expression. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have

been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

5 Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding APOP.

10 Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of
15 candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis.

20 Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA sequences encoding APOP. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible
25 modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine,
30 guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

An additional embodiment of the invention encompasses a method for screening for a compound which is effective in altering expression of a polynucleotide encoding APOP. Compounds which may be effective in altering expression of a specific polynucleotide may include, but are not limited to, oligonucleotides, antisense oligonucleotides, triple helix-forming oligonucleotides,
35 transcription factors and other polypeptide transcriptional regulators, and non-macromolecular

chemical entities which are capable of interacting with specific polynucleotide sequences. Effective compounds may alter polynucleotide expression by acting as either inhibitors or promoters of polynucleotide expression. Thus, in the treatment of disorders associated with increased APOP expression or activity, a compound which specifically inhibits expression of the polynucleotide
5 encoding APOP may be therapeutically useful, and in the treatment of disorders associated with decreased APOP expression or activity, a compound which specifically promotes expression of the polynucleotide encoding APOP may be therapeutically useful.

At least one, and up to a plurality, of test compounds may be screened for effectiveness in altering expression of a specific polynucleotide. A test compound may be obtained by any method
10 commonly known in the art, including chemical modification of a compound known to be effective in altering polynucleotide expression; selection from an existing, commercially-available or proprietary library of naturally-occurring or non-natural chemical compounds; rational design of a compound based on chemical and/or structural properties of the target polynucleotide; and selection from a library of chemical compounds created combinatorially or randomly. A sample comprising a
15 polynucleotide encoding APOP is exposed to at least one test compound thus obtained. The sample may comprise, for example, an intact or permeabilized cell, or an *in vitro* cell-free or reconstituted biochemical system. Alterations in the expression of a polynucleotide encoding APOP are assayed by any method commonly known in the art. Typically, the expression of a specific nucleotide is detected by hybridization with a probe having a nucleotide sequence complementary to the sequence of the
20 polynucleotide encoding APOP. The amount of hybridization may be quantified, thus forming the basis for a comparison of the expression of the polynucleotide both with and without exposure to one or more test compounds. Detection of a change in the expression of a polynucleotide exposed to a test compound indicates that the test compound is effective in altering the expression of the polynucleotide. A screen for a compound effective in altering expression of a specific polynucleotide
25 can be carried out, for example, using a *Schizosaccharomyces pombe* gene expression system (Atkins, D. et al. (1999) U.S. Patent No. 5,932,435; Arndt, G.M. et al. (2000) Nucleic Acids Res. 28:E15) or a human cell line such as HeLa cell (Clarke, M.L. et al. (2000) Biochem. Biophys. Res. Commun. 268:8-13). A particular embodiment of the present invention involves screening a combinatorial library of oligonucleotides (such as deoxyribonucleotides, ribonucleotides, peptide nucleic acids, and
30 modified oligonucleotides) for antisense activity against a specific polynucleotide sequence (Bruce, T.W. et al. (1997) U.S. Patent No. 5,686,242; Bruce, T.W. et al. (2000) U.S. Patent No. 6,022,691).

Many methods for introducing vectors into cells or tissues are available and equally suitable for use *in vivo*, *in vitro*, and *ex vivo*. For *ex vivo* therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient.
35 Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved

using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and
5 monkeys.

An additional embodiment of the invention relates to the administration of a composition which generally comprises an active ingredient formulated with a pharmaceutically acceptable excipient. Excipients may include, for example, sugars, starches, celluloses, gums, and proteins. Various formulations are commonly known and are thoroughly discussed in the latest edition of Remington's
10 Pharmaceutical Sciences (Maack Publishing, Easton PA). Such compositions may consist of APOP, antibodies to APOP, and mimetics, agonists, antagonists, or inhibitors of APOP.

The compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, pulmonary, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical,
15 sublingual, or rectal means.

Compositions for pulmonary administration may be prepared in liquid or dry powder form. These compositions are generally aerosolized immediately prior to inhalation by the patient. In the case of small molecules (e.g. traditional low molecular weight organic drugs), aerosol delivery of fast-acting formulations is well-known in the art. In the case of macromolecules (e.g. larger peptides and proteins),
20 recent developments in the field of pulmonary delivery via the alveolar region of the lung have enabled the practical delivery of drugs such as insulin to blood circulation (see, e.g., Patton, J.S. et al., U.S. Patent No. 5,997,848). Pulmonary delivery has the advantage of administration without needle injection, and obviates the need for potentially toxic penetration enhancers.

Compositions suitable for use in the invention include compositions wherein the active
25 ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

Specialized forms of compositions may be prepared for direct intracellular delivery of macromolecules comprising APOP or fragments thereof. For example, liposome preparations containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of the
30 macromolecule. Alternatively, APOP or a fragment thereof may be joined to a short cationic N-terminal portion from the HIV Tat-1 protein. Fusion proteins thus generated have been found to transduce into the cells of all tissues, including the brain, in a mouse model system (Schwarze, S.R. et al. (1999) Science 285:1569-1572).

For any compound, the therapeutically effective dose can be estimated initially either in cell

culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, monkeys, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

- 5 A therapeutically effective dose refers to that amount of active ingredient, for example APOP or fragments thereof, antibodies of APOP, and agonists, antagonists or inhibitors of APOP, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED_{50} (the dose therapeutically effective in 50% of the population) or LD_{50} (the dose
- 10 lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD_{50}/ED_{50} ratio. Compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity.
- 15 The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

 The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity

20 of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

 Normal dosage amounts may vary from about 0.1 μg to 100,000 μg , up to a total dose of

25 about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

30 **DIAGNOSTICS**

 In another embodiment, antibodies which specifically bind APOP may be used for the diagnosis of disorders characterized by expression of APOP, or in assays to monitor patients being treated with APOP or agonists, antagonists, or inhibitors of APOP. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for APOP

include methods which utilize the antibody and a label to detect APOP in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

5 A variety of protocols for measuring APOP, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of APOP expression. Normal or standard values for APOP expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibody to APOP under conditions suitable for complex formation. The amount of standard complex formation may be
10 quantitated by various methods, such as photometric means. Quantities of APOP expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding APOP may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences,
15 complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify gene expression in biopsied tissues in which expression of APOP may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of APOP, and to monitor regulation of APOP levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide
20 sequences, including genomic sequences, encoding APOP or closely related molecules may be used to identify nucleic acid sequences which encode APOP. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the probe identifies only naturally occurring sequences encoding APOP, allelic variants, or related
25 sequences.

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the APOP encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:6-10 or from genomic sequences including promoters, enhancers, and introns of the APOP gene.

30 Means for producing specific hybridization probes for DNAs encoding APOP include the cloning of polynucleotide sequences encoding APOP or APOP derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety

of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S , or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding APOP may be used for the diagnosis of disorders associated with expression of APOP. Examples of such disorders include, but are not limited to, an immune disorder such as inflammation, actinic keratosis, acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, arteriosclerosis, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, bronchitis, bursitis, cholecystitis, cirrhosis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, paroxysmal nocturnal hemoglobinuria, hepatitis, hypereosinophilia, irritable bowel syndrome, episodic lymphopenia with lymphocytotoxins, mixed connective tissue disease (MCTD), multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, myelofibrosis, osteoarthritis, osteoporosis, pancreatitis, polycythemia vera, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, primary thrombocythemia, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, trauma, and hematopoietic cancer including lymphoma, leukemia, and myeloma; and a cell proliferative disorder such as actinic keratosis, paroxysmal nocturnal hemoglobinuria, and cancers including adenocarcinoma, melanoma, sarcoma, teratocarcinoma, and, in particular, a cancer of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus. The polynucleotide sequences encoding APOP may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered APOP expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding APOP may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding APOP may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding APOP in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy

of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of APOP, a normal or standard profile for expression is established. This may be accomplished by combining
5 body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding APOP, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from
10 patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from
15 successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual
20 clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding APOP may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically,
25 or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide encoding APOP, or a fragment of a polynucleotide complementary to the polynucleotide encoding APOP, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

30 In a particular aspect, oligonucleotide primers derived from the polynucleotide sequences encoding APOP may be used to detect single nucleotide polymorphisms (SNPs). SNPs are substitutions, insertions and deletions that are a frequent cause of inherited or acquired genetic disease in humans. Methods of SNP detection include, but are not limited to, single-stranded conformation polymorphism (SSCP) and fluorescent SSCP (fSSCP) methods. In SSCP, oligonucleotide primers

derived from the polynucleotide sequences encoding APOP are used to amplify DNA using the polymerase chain reaction (PCR). The DNA may be derived, for example, from diseased or normal tissue, biopsy samples, bodily fluids, and the like. SNPs in the DNA cause differences in the secondary and tertiary structures of PCR products in single-stranded form, and these differences are detectable using gel electrophoresis in non-denaturing gels. In fSCCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplimers in high-throughput equipment such as DNA sequencing machines. Additionally, sequence database analysis methods, termed in silico SNP (isSNP), are capable of identifying polymorphisms by comparing the sequence of individual overlapping DNA fragments which assemble into a common consensus sequence. These computer-based methods filter out sequence variations due to laboratory preparation of DNA and sequencing errors using statistical models and automated analyses of DNA sequence chromatograms. In the alternative, SNPs may be detected and characterized by mass spectrometry using, for example, the high throughput MASSARRAY system (Sequenom, Inc., San Diego CA).

Methods which may also be used to quantify the expression of APOP include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer or polynucleotide of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as elements on a microarray. The microarray can be used in transcript imaging techniques which monitor the relative expression levels of large numbers of genes simultaneously as described in Seilhamer, J.J. et al., "Comparative Gene Transcript Analysis," U.S. Patent No. 5,840,484, incorporated herein by reference. The microarray may also be used to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, to monitor progression/regression of disease as a function of gene expression, and to develop and monitor the activities of therapeutic agents in the treatment of disease. In particular, this information may be used to develop a pharmacogenomic profile of a patient in order to select the most appropriate and effective treatment regimen for that patient. For example, therapeutic agents which are highly effective and display the fewest side effects may be selected for a patient based on his/her pharmacogenomic profile.

In another embodiment, antibodies specific for APOP, or APOP or fragments thereof may be used as elements on a microarray. The microarray may be used to monitor or measure protein-protein

interactions, drug-target interactions, and gene expression profiles, as described above.

A particular embodiment relates to the use of the polynucleotides of the present invention to generate a transcript image of a tissue or cell type. A transcript image represents the global pattern of gene expression by a particular tissue or cell type. Global gene expression patterns are analyzed by
5 quantifying the number of expressed genes and their relative abundance under given conditions and at a given time. (See Seilhamer et al., "Comparative Gene Transcript Analysis," U.S. Patent Number 5,840,484, expressly incorporated by reference herein.) Thus a transcript image may be generated by hybridizing the polynucleotides of the present invention or their complements to the totality of transcripts or reverse transcripts of a particular tissue or cell type. In one embodiment, the
10 hybridization takes place in high-throughput format, wherein the polynucleotides of the present invention or their complements comprise a subset of a plurality of elements on a microarray. The resultant transcript image would provide a profile of gene activity.

Transcript images may be generated using transcripts isolated from tissues, cell lines, biopsies, or other biological samples. The transcript image may thus reflect gene expression in vivo, as in the
15 case of a tissue or biopsy sample, or in vitro, as in the case of a cell line.

Transcript images which profile the expression of the polynucleotides of the present invention may also be used in conjunction with in vitro model systems and preclinical evaluation of pharmaceuticals, as well as toxicological testing of industrial and naturally-occurring environmental compounds. All compounds induce characteristic gene expression patterns, frequently termed
20 molecular fingerprints or toxicant signatures, which are indicative of mechanisms of action and toxicity (Nuwaysir, E.F. et al. (1999) Mol. Carcinog. 24:153-159; Steiner, S. and N.L. Anderson (2000) Toxicol. Lett. 112-113:467-471, expressly incorporated by reference herein). If a test compound has a signature similar to that of a compound with known toxicity, it is likely to share those toxic properties. These fingerprints or signatures are most useful and refined when they contain expression information
25 from a large number of genes and gene families. Ideally, a genome-wide measurement of expression provides the highest quality signature. Even genes whose expression is not altered by any tested compounds are important as well, as the levels of expression of these genes are used to normalize the rest of the expression data. The normalization procedure is useful for comparison of expression data after treatment with different compounds. While the assignment of gene function to elements of a
30 toxicant signature aids in interpretation of toxicity mechanisms, knowledge of gene function is not necessary for the statistical matching of signatures which leads to prediction of toxicity. (See, for example, Press Release 00-02 from the National Institute of Environmental Health Sciences, released February 29, 2000, available at <http://www.niehs.nih.gov/oc/news/toxchip.htm>.) Therefore, it is important and desirable in toxicological screening using toxicant signatures to include all expressed

gene sequences.

In one embodiment, the toxicity of a test compound is assessed by treating a biological sample containing nucleic acids with the test compound. Nucleic acids that are expressed in the treated biological sample are hybridized with one or more probes specific to the polynucleotides of the present invention, so that transcript levels corresponding to the polynucleotides of the present invention may be quantified. The transcript levels in the treated biological sample are compared with levels in an untreated biological sample. Differences in the transcript levels between the two samples are indicative of a toxic response caused by the test compound in the treated sample.

Another particular embodiment relates to the use of the polypeptide sequences of the present invention to analyze the proteome of a tissue or cell type. The term proteome refers to the global pattern of protein expression in a particular tissue or cell type. Each protein component of a proteome can be subjected individually to further analysis. Proteome expression patterns, or profiles, are analyzed by quantifying the number of expressed proteins and their relative abundance under given conditions and at a given time. A profile of a cell's proteome may thus be generated by separating and analyzing the polypeptides of a particular tissue or cell type. In one embodiment, the separation is achieved using two-dimensional gel electrophoresis, in which proteins from a sample are separated by isoelectric focusing in the first dimension, and then according to molecular weight by sodium dodecyl sulfate slab gel electrophoresis in the second dimension (Steiner and Anderson, supra). The proteins are visualized in the gel as discrete and uniquely positioned spots, typically by staining the gel with an agent such as Coomassie Blue or silver or fluorescent stains. The optical density of each protein spot is generally proportional to the level of the protein in the sample. The optical densities of equivalently positioned protein spots from different samples, for example, from biological samples either treated or untreated with a test compound or therapeutic agent, are compared to identify any changes in protein spot density related to the treatment. The proteins in the spots are partially sequenced using, for example, standard methods employing chemical or enzymatic cleavage followed by mass spectrometry. The identity of the protein in a spot may be determined by comparing its partial sequence, preferably of at least 5 contiguous amino acid residues, to the polypeptide sequences of the present invention. In some cases, further sequence data may be obtained for definitive protein identification.

A proteomic profile may also be generated using antibodies specific for APOP to quantify the levels of APOP expression. In one embodiment, the antibodies are used as elements on a microarray, and protein expression levels are quantified by exposing the microarray to the sample and detecting the levels of protein bound to each array element (Lueking, A. et al. (1999) Anal. Biochem. 270:103-111; Mendoze, L.G. et al. (1999) Biotechniques 27:778-788). Detection may be performed by a variety of methods known in the art, for example, by reacting the proteins in the sample with a thiol- or amino-reactive fluorescent compound and detecting the amount of fluorescence bound at each array element.

Toxicant signatures at the proteome level are also useful for toxicological screening, and should be analyzed in parallel with toxicant signatures at the transcript level. There is a poor correlation between transcript and protein abundances for some proteins in some tissues (Anderson, N.L. and J. Seilhamer (1997) *Electrophoresis* 18:533-537), so proteome toxicant signatures may be useful in the analysis of compounds which do not significantly affect the transcript image, but which alter the proteomic profile. In addition, the analysis of transcripts in body fluids is difficult, due to rapid degradation of mRNA, so proteomic profiling may be more reliable and informative in such cases.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins that are expressed in the treated biological sample are separated so that the amount of each protein can be quantified. The amount of each protein is compared to the amount of the corresponding protein in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample. Individual proteins are identified by sequencing the amino acid residues of the individual proteins and comparing these partial sequences to the polypeptides of the present invention.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins from the biological sample are incubated with antibodies specific to the polypeptides of the present invention. The amount of protein recognized by the antibodies is quantified. The amount of protein in the treated biological sample is compared with the amount in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) *Proc. Natl. Acad. Sci. USA* 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) *Proc. Natl. Acad. Sci. USA* 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.) Various types of microarrays are well known and thoroughly described in DNA Microarrays: A Practical Approach, M. Schena, ed. (1999) Oxford University Press, London, hereby expressly incorporated by reference.

In another embodiment of the invention, nucleic acid sequences encoding APOP may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. Either coding or noncoding sequences may be used, and in some instances, noncoding sequences may be preferable over coding sequences. For example, conservation of a coding sequence among members of a multi-gene family may potentially cause undesired cross hybridization during chromosomal mapping. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1

constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.) Once mapped, the nucleic acid sequences of the invention may be used to develop genetic linkage maps, for example, which correlate the inheritance of a disease state with the inheritance of a particular chromosome region or restriction fragment length polymorphism (RFLP). (See, e.g.,
5 Lander, E.S. and D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM)
10 World Wide Web site. Correlation between the location of the gene encoding APOP on a physical map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder and thus may further positional cloning efforts.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps.
15 Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the exact chromosomal locus is not known. This information is valuable to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the gene or genes responsible for a disease or syndrome have been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences
20 mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the instant invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, APOP, its catalytic or immunogenic fragments, or
25 oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between APOP and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds
30 having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with APOP, or fragments thereof, and washed. Bound APOP is then detected by methods well known in the art. Purified APOP can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively,
35 non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding APOP specifically compete with a test compound for binding APOP. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with APOP.

5 In additional embodiments, the nucleotide sequences which encode APOP may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding
10 description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications and publications, mentioned above and below, in particular U.S. Ser. No. 60/153,403, are hereby expressly incorporated by reference.

15

EXAMPLES

I. Construction of cDNA Libraries

RNA was purchased from Clontech or isolated from tissues described in Table 4. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed
20 in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA
25 purity. In some cases, RNA was treated with DNase. For most libraries, poly(A+) RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

30 In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERScript plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, supra, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic
35 oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the

appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g.,

- 5 PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), pcDNA2.1 plasmid (Invitrogen, Carlsbad CA), or pINCY plasmid (Incyte Genomics, Palo Alto CA). Recombinant plasmids were transformed into competent *E. coli* cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5 α , DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

- 10 Plasmids obtained as described in Example I were recovered from host cells by in vivo excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid
15 purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

- Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-
20 well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

- Incyte cDNA recovered in plasmids as described in Example II were sequenced as follows.
25 Sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (PE Biosystems) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI
30 PRISM BIGDYE Terminator cycle sequencing ready reaction kit (PE Biosystems). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (PE Biosystems) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA

sequences were identified using standard methods (reviewed in Ausubel, 1997, supra, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VI.

The polynucleotide sequences derived from cDNA sequencing were assembled and analyzed using a combination of software programs which utilize algorithms well known to those skilled in the art. Table 5 summarizes the tools, programs, and algorithms used and provides applicable descriptions, references, and threshold parameters. The first column of Table 5 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score, the greater the homology between two sequences). Sequences were analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments were generated using the default parameters specified by the clustal algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

The polynucleotide sequences were validated by removing vector, linker, and polyA sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The sequences were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM, and PFAM to acquire annotation using programs based on BLAST, FASTA, and BLIMPS. The sequences were assembled into full length polynucleotide sequences using programs based on Phred, Phrap, and Consed, and were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length amino acid sequences, and these full length sequences were subsequently analyzed by querying against databases such as the GenBank databases (described above), SwissProt, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and Hidden Markov Model (HMM)-based protein family databases such as PFAM. HMM is a probabilistic approach which analyzes consensus primary structures of gene families. (See, e.g., Eddy, S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365.)

The programs described above for the assembly and analysis of full length polynucleotide and amino acid sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:6-10. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies were described in The Invention section above.

IV. Analysis of Polynucleotide Expression

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, supra, ch. 7; Ausubel, 1995, supra, ch. 4 and 16.)

5 Analogous computer techniques applying BLAST were used to search for identical or related molecules in cDNA databases such as GenBank or LIFESEQ (Incyte Genomics). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$10 \quad \frac{\text{BLAST Score} \times \text{Percent Identity}}{5 \times \text{minimum \{length(Seq. 1), length(Seq. 2)\}}}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. The product score is a normalized value between 0 and 100, and is calculated as follows: the BLAST score is multiplied by the percent nucleotide identity and the product is divided by (5 times the length of the shorter of the two sequences). The BLAST score is calculated by assigning a score of +5 for every base that matches in a high-scoring segment pair (HSP), and -4 for every mismatch. Two sequences may share more than one HSP (separated by gaps). If there is more than one HSP, then the pair with the highest BLAST score is used to calculate the product score. The product score represents a balance between fractional overlap and quality in a BLAST alignment. For example, a product score of 100 is produced only for 100% identity over the entire length of the shorter of the two sequences being compared. A product score of 70 is produced either by 100% identity and 70% overlap at one end, or by 88% identity and 100% overlap at the other. A product score of 50 is produced either by 100% identity and 50% overlap at one end, or 79% identity and 100% overlap.

25 The results of northern analyses are reported as a percentage distribution of libraries in which the transcript encoding APOP occurred. Analysis involved the categorization of cDNA libraries by organ/tissue and disease. The organ/tissue categories included cardiovascular, dermatologic, developmental, endocrine, gastrointestinal, hematopoietic/immune, musculoskeletal, nervous, reproductive, and urologic. The disease/condition categories included cancer, inflammation, trauma, cell proliferation, neurological, and pooled. For each category, the number of libraries expressing the sequence of interest was counted and divided by the total number of libraries across all categories. Percentage values of tissue-specific and disease- or condition-specific expression are reported in Table 3.

30 V. Chromosomal Mapping of APOP Encoding Polynucleotides

The cDNA sequences which were used to assemble SEQ ID NO:6-10 were compared with sequences from the Incyte LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:6-10 were assembled into clusters of contiguous and overlapping sequences using
5 assembly algorithms such as Phrap (Table 5). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and Généthon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO., to that map location.

10 The genetic map locations of SEQ ID NO:6 and SEQ ID NO:8 are described in The Invention as ranges, or intervals, of human chromosomes. The map position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-arm. The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in humans, although this can vary widely (due to
15 hot and cold spots of recombination.) The cM distances are based on genetic markers mapped by Généthon which provide boundaries for radiation hybrid markers whose sequences were included in each of the clusters.

VI. Extension of APOP Encoding Polynucleotides

The full length nucleic acid sequences of SEQ ID NO:6-10 were produced by extension of an
20 appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer, to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at
25 temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR
30 was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+} , $(NH_4)_2SO_4$, and β -mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C,
35 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the

alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 µl PICOGREEN
5 quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE
and 0.5 µl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar,
Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoröskan II
(Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the
concentration of DNA. A 5 µl to 10 µl aliquot of the reaction mixture was analyzed by electrophoresis
10 on a 1 % agarose mini-gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates,
digested with CviII cholera virus endonuclease (Molecular Biology Research, Madison WI), and
sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For
shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose
15 gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were
religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham
Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs,
and transfected into competent *E. coli* cells. Transformed cells were selected on antibiotic-containing
media, and individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x
20 carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham
Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1:
94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4
repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN
25 reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified
using the same conditions as described above. Samples were diluted with 20% dimethylsulfoxide (1:2,
v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC
DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle
sequencing ready reaction kit (PE Biosystems).

30 In like manner, the polynucleotide sequences of SEQ ID NO:6-10 are used to obtain 5'
regulatory sequences using the procedure above, along with oligonucleotides designed for such
extension, and an appropriate genomic library.

VII. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:6-10 are employed to screen cDNAs, genomic

DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments.

Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μ Ci of [γ - 32 P] adenosine

- 5 triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or
- 10 Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate.

- 15 Hybridization patterns are visualized using autoradiography or an alternative imaging means and compared.

VIII. Microarrays

The linkage or synthesis of array elements upon a microarray can be achieved utilizing photolithography, piezoelectric printing (ink-jet printing, See, e.g., Baldeschweiler, *supra*), mechanical

20 microspotting technologies, and derivatives thereof. The substrate in each of the aforementioned technologies should be uniform and solid with a non-porous surface (Skena (1999), *supra*). Suggested substrates include silicon, silica, glass slides, glass chips, and silicon wafers. Alternatively, a procedure analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be

25 produced using available methods and machines well known to those of ordinary skill in the art and may contain any appropriate number of elements. (See, e.g., Skena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645; Marshall, A. and J. Hodgson (1998) Nat. Biotechnol. 16:27-31.)

- Full length cDNAs, Expressed Sequence Tags (ESTs), or fragments or oligomers thereof may
- 30 comprise the elements of the microarray. Fragments or oligomers suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). The array elements are hybridized with polynucleotides in a biological sample. The polynucleotides in the biological sample are conjugated to a fluorescent label or other molecular tag for ease of detection. After hybridization, nonhybridized nucleotides from the biological sample are removed, and a

fluorescence scanner is used to detect hybridization at each array element. Alternatively, laser desorption and mass spectrometry may be used for detection of hybridization. The degree of complementarity and the relative abundance of each polynucleotide which hybridizes to an element on the microarray may be assessed. In one embodiment, microarray preparation and usage is described in detail below.

Tissue or Cell Sample Preparation

Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and poly(A)⁺ RNA is purified using the oligo-(dT) cellulose method. Each poly(A)⁺ RNA sample is reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/ μ l oligo-(dT) primer (21mer), 1X first strand buffer, 0.03 units/ μ l RNase inhibitor, 500 μ M dATP, 500 μ M dGTP, 500 μ M dTTP, 40 μ M dCTP, 40 μ M dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Pharmacia Biotech). The reverse transcription reaction is performed in a 25 ml volume containing 200 ng poly(A)⁺ RNA with GEMBRIGHT kits (Incyte). Specific control poly(A)⁺ RNAs are synthesized by *in vitro* transcription from non-coding yeast genomic DNA. After incubation at 37 °C for 2 hr, each reaction sample (one with Cy3 and another with Cy5 labeling) is treated with 2.5 ml of 0.5M sodium hydroxide and incubated for 20 minutes at 85 °C to stop the reaction and degrade the RNA. Samples are purified using two successive CHROMA SPIN 30 gel filtration spin columns (CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA) and after combining, both reaction samples are ethanol precipitated using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The sample is then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and resuspended in 14 μ l 5X SSC/0.2% SDS.

Microarray Preparation

Sequences of the present invention are used to generate array elements. Each array element is amplified from bacterial cells containing vectors with cloned cDNA inserts. PCR amplification uses primers complementary to the vector sequences flanking the cDNA insert. Array elements are amplified in thirty cycles of PCR from an initial quantity of 1-2 ng to a final quantity greater than 5 μ g. Amplified array elements are then purified using SEPHACRYL-400 (Amersham Pharmacia Biotech).

Purified array elements are immobilized on polymer-coated glass slides. Glass microscope slides (Corning) are cleaned by ultrasound in 0.1% SDS and acetone, with extensive distilled water washes between and after treatments. Glass slides are etched in 4% hydrofluoric acid (VWR Scientific Products Corporation (VWR), West Chester PA), washed extensively in distilled water, and coated with 0.05% aminopropyl silane (Sigma) in 95% ethanol. Coated slides are cured in a 110°C oven.

Array elements are applied to the coated glass substrate using a procedure described in US

Patent No. 5,807,522, incorporated herein by reference. 1 μ l of the array element DNA, at an average concentration of 100 ng/ μ l, is loaded into the open capillary printing element by a high-speed robotic apparatus. The apparatus then deposits about 5 nl of array element sample per slide.

Microarrays are UV-crosslinked using a STRATALINKER UV-crosslinker (Stratagene).

- 5 Microarrays are washed at room temperature once in 0.2% SDS and three times in distilled water. Non-specific binding sites are blocked by incubation of microarrays in 0.2% casein in phosphate buffered saline (PBS) (Tropix, Inc., Bedford MA) for 30 minutes at 60 °C followed by washes in 0.2% SDS and distilled water as before.

Hybridization

- 10 Hybridization reactions contain 9 μ l of sample mixture consisting of 0.2 μ g each of Cy3 and Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The sample mixture is heated to 65 °C for 5 minutes and is aliquoted onto the microarray surface and covered with an 1.8 cm² coverslip. The arrays are transferred to a waterproof chamber having a cavity just slightly larger than a microscope slide. The chamber is kept at 100% humidity internally by the addition of
- 15 140 μ l of 5X SSC in a corner of the chamber. The chamber containing the arrays is incubated for about 6.5 hours at 60 °C. The arrays are washed for 10 min at 45 °C in a first wash buffer (1X SSC, 0.1% SDS), three times for 10 minutes each at 45 °C in a second wash buffer (0.1X SSC), and dried.

Detection

- Reporter-labeled hybridization complexes are detected with a microscope equipped with an
- 20 Innova 70 mixed gas 10 W laser (Coherent, Inc., Santa Clara CA) capable of generating spectral lines at 488 nm for excitation of Cy3 and at 632 nm for excitation of Cy5. The excitation laser light is focused on the array using a 20X microscope objective (Nikon, Inc., Melville NY). The slide containing the array is placed on a computer-controlled X-Y stage on the microscope and raster-scanned past the objective. The 1.8 cm x 1.8 cm array used in the present example is scanned with a
- 25 resolution of 20 micrometers.

- In two separate scans, a mixed gas multiline laser excites the two fluorophores sequentially. Emitted light is split, based on wavelength, into two photomultiplier tube detectors (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) corresponding to the two fluorophores. Appropriate filters positioned between the array and the photomultiplier tubes are used to filter the signals. The
- 30 emission maxima of the fluorophores used are 565 nm for Cy3 and 650 nm for Cy5. Each array is typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source, although the apparatus is capable of recording the spectra from both fluorophores simultaneously.

- The sensitivity of the scans is typically calibrated using the signal intensity generated by a cDNA control species added to the sample mixture at a known concentration. A specific location on
- 35 the array contains a complementary DNA sequence, allowing the intensity of the signal at that

location to be correlated with a weight ratio of hybridizing species of 1:100,000. When two samples from different sources (e.g., representing test and control cells), each labeled with a different fluorophore, are hybridized to a single array for the purpose of identifying genes that are differentially expressed, the calibration is done by labeling samples of the calibrating cDNA with the two
5 fluorophores and adding identical amounts of each to the hybridization mixture.

The output of the photomultiplier tube is digitized using a 12-bit RTI-835H analog-to-digital (A/D) conversion board (Analog Devices, Inc., Norwood MA) installed in an IBM-compatible PC computer. The digitized data are displayed as an image where the signal intensity is mapped using a linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high
10 signal). The data is also analyzed quantitatively. Where two different fluorophores are excited and measured simultaneously, the data are first corrected for optical crosstalk (due to overlapping emission spectra) between the fluorophores using each fluorophore's emission spectrum.

A grid is superimposed over the fluorescence signal image such that the signal from each spot is centered in each element of the grid. The fluorescence signal within each element is then integrated
15 to obtain a numerical value corresponding to the average intensity of the signal. The software used for signal analysis is the GEMTOOLS gene expression analysis program (Incyte).

IX. Complementary Polynucleotides

Sequences complementary to the APOP-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring APOP. Although use of oligonucleotides
20 comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of APOP. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is
25 designed to prevent ribosomal binding to the APOP-encoding transcript.

X. Expression of APOP

Expression and purification of APOP is achieved using bacterial or virus-based expression systems. For expression of APOP in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription.
30 Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express APOP upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of APOP in eukaryotic cells is achieved by infecting insect or mammalian cell lines with
35 recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as

baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding APOP by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, APOP is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from APOP at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, *supra*, ch. 10 and 16). Purified APOP obtained by these methods can be used directly in the assays shown in Examples XI and XIV.

XI. Demonstration of APOP Activity

An assay for APOP activity measures the induction of apoptosis when APOP is expressed at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include pCMV SPORT (Life Technologies, Gaithersburg MD) and pCR 3.1 (Invitrogen, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10 μ g of recombinant vector are transiently transfected into a human cell line, preferably of endothelial or hematopoietic origin, using either liposome formulations or electroporation. 1-2 μ g of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP) (Clontech, Palo Alto CA), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate their apoptotic state. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These

events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with
5 specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of APOP on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding APOP and either CD64 or CD64-GFP. CD64 and CD64-
10 GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding APOP and other genes of interest can be analyzed by northern analysis or
15 microarray techniques.

XII. Production of APOP Specific Antibodies

APOP substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) *Methods Enzymol.* 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

20 Alternatively, the APOP amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

25 Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (PE Biosystems) using Fmoc chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-APOP
30 activity by, for example, binding the peptide or APOP to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XIII. Purification of Naturally Occurring APOP Using Specific Antibodies

Naturally occurring or recombinant APOP is substantially purified by immunoaffinity chromatography using antibodies specific for APOP. An immunoaffinity column is constructed by

covalently coupling anti-APOP antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing APOP are passed over the immunoaffinity column, and the column is washed
5 under conditions that allow the preferential absorbance of APOP (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/APOP binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and APOP is collected.

XIV. Identification of Molecules Which Interact with APOP

10 APOP, or biologically active fragments thereof, are labeled with ^{125}I Bolton-Hunter reagent. (See, e.g., Bolton A.E. and W.M. Hunter (1973) *Biochem. J.* 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled APOP, washed, and any wells with labeled APOP complex are assayed. Data obtained using different concentrations of APOP are used to calculate values for the number, affinity, and association of APOP with the candidate
15 molecules.

Alternatively, molecules interacting with APOP are analyzed using the yeast two-hybrid system as described in Fields, S. and O. Song (1989, *Nature* 340:245-246), or using commercially available kits based on the two-hybrid system, such as the MATCHMAKER system (Clontech).

APOP may also be used in the PATHCALLING process (CuraGen Corp., New Haven CT)
20 which employs the yeast two-hybrid system in a high-throughput manner to determine all interactions between the proteins encoded by two large libraries of genes (Nandabalan, K. et al. (2000) U.S. Patent No. 6,057,101).

Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention.
25 Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

Table 1

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
1	6	898712	BRSTTUT03	377292F1 (NEUTFMT01), 898632R1 (BRSTTUT03), 898712H1 (BRSTTUT03), 1616481F6 (BRAITUT12)
2	7	1301667	BRSTNOT07	1216205R1 (BRSTTUT01), 1301667H1 (BRSTNOT07), 1855956F6 (PROSNOT18), 1864511F6 (PROSNOT19), 1890039F6 (BLADTUT07), 2220053F6 (LUNGNOT18), 3950887F6 (DRGCNOT01)
3	8	1623420	BRAITUT13	060307F1 (LUNGNOT01), 1492891T1 (PROSNON01), 1781466R6 (PGANNON02), 2375493F6 (ISLTNOT01)
4	9	2240612	PANCTUT02	595459H1 (BRAVUNT02), 941628T6 (ADRENOT03), 2240612H1 (PANCTUT02), 2509190T6 (CONUTUT01), 2623037R6 (KERANOT02), 2743148F6 (BRSTTUT14), SZAU000083V1
5	10	4270927	ENDCNON02	2047528H1 (ENDCNOT02), 4270927H1 (ENDCNON02), 4270927T6 (ENDCNON02)

Table 2

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods & Databases
1	115	S50		Signal peptide: M1-L31 Zinc finger: C67-C101	Apoptosis inhibitor (g3822314) <u>Lymantria dispar nucleopolyhedrovirus</u>	BLAST-GenBank SPScan HMMER-PFAM MOTIFS
2	311	T215 S264 S269 S273 S274 S275 S286	N213	Transmembrane domains: L190-F209, S228-W246 Homeobox domain: V12-P54 Protein transmembrane longevity-assurance protein 1: F45-C288	Similar to S. cerevisiae longevity-assurance protein 1 (g1123105) <u>Caenorhabditis elegans</u>	BLAST-GenBank HMMER BLIMPS-PRINTS BLIMPS-BLOCKS BLAST-PRODOM BLAST-DOMO MOTIFS
3	219	S42 T43 S49 S56 T57 T63 T64 S105 S155 S199 S7 T87 S171		Antigen melanoma-associated multigene family tumor polymorphism Mage-1: N83-E204	Apoptosis related protein APR-1 (g4982472) <u>Homo sapiens</u>	BLAST-GenBank BLAST-PRODOM BLAST-DOMO MOTIFS
4	358	S132 S152 T340 S92 T159 Y251		Signal peptide: M1-G19 Protein zinc finger RP-8 apoptosis zinc-finger: H176-D353	PDCD2 (programmed cell death-2/RP8 homolog) (g3790133) <u>Homo sapiens</u>	BLAST-GenBank SPScan BLAST-PRODOM BLAST-DOMO MOTIFS
5	269	S154 T6 S28 S40 S112 S135 T220 T227 S256 T34 S35 S57 S161 Y44 Y169	N49		Apoptosis repressor ARC (g3126907) <u>Homo sapiens</u>	BLAST-GenBank

Table 3

Nucleotide SEQ ID NO:	Selected Fragments	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
6	195-239 240-299 300-359 360-419 420-479 480-539	Nervous (0.263) Reproductive (0.263)	Inflammation/Trauma (0.474) Cancer (0.421) Cell Proliferation (0.105)	PSPORT1
7	480-524	Reproductive (0.349) Nervous (0.143) Hematopoietic/Immune (0.127)	Cancer (0.635) Inflammation/Trauma (0.265)	pINCY
8	186-230	Nervous (0.278) Reproductive (0.215) Gastrointestinal (0.139)	Cancer (0.418) Inflammation/Trauma (0.240) Cell Proliferation (0.177)	pINCY
9	375-419	Gastrointestinal (0.294) Nervous (0.235) Reproductive (0.118)	Cancer (0.470) Cell Proliferation (0.240) Inflammation/Trauma (0.236)	pINCY
10	6-50 51-350 351-650 651-812	Cardiovascular (0.400) Reproductive (0.400) Dermatologic (0.200)	Cell Proliferation (0.600) Cancer (0.400)	pINCY

Table 4

SEQ ID NO:	Library	Library Comment
6	BRSTTUT03	Library was constructed using RNA isolated from breast tumor tissue removed from a 58-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology indicated multicentric invasive grade 4 lobular carcinoma. The mass was identified in the upper outer quadrant, and three separate nodules were found in the lower outer quadrant of the left breast. Patient history included skin cancer, rheumatic heart disease, osteoarthritis, and tuberculosis. Family history included cerebrovascular disease, coronary artery aneurysm, breast cancer, prostate cancer, atherosclerotic coronary artery disease, and type I diabetes.
7	BRSTNOT07	Library was constructed using RNA isolated from diseased breast tissue removed from a 43-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology indicated mildly proliferative fibrocystic changes with epithelial hyperplasia, papillomatosis, and duct ectasia. Pathology for the associated tumor tissue indicated invasive grade 4, nuclear grade 3 mammary adenocarcinoma with extensive comedo necrosis. Family history included epilepsy, cardiovascular disease, and type II diabetes.
8	BRAITUT13	Library was constructed using RNA isolated from brain tumor tissue removed from the left frontal lobe of a 68-year-old Caucasian male during excision of a cerebral meningeal lesion. Pathology indicated a meningioma in the left frontal lobe.
9	PANCTUT02	Library was constructed using RNA isolated from pancreatic tumor tissue removed from a 45-year-old Caucasian female during radical pancreaticoduodenectomy. Pathology indicated a grade 4 anaplastic carcinoma. Family history included benign hypertension, hyperlipidemia and atherosclerotic coronary artery disease.
10	ENDCNON02	This normalized coronary artery endothelial cell tissue library was constructed from 444,000 independent clones from ENDCNOT04, an endothelial cell library. Starting RNA was isolated from coronary artery endothelial cell tissue removed from a 3-year-old Caucasian male. This library was normalized in two rounds using conditions adapted from Soares et al. (1994; Proc. Natl. Acad. Sci. USA 91:9228-9232) and Bonaldo et al. (1996; Genome Res. 6:791-806), using a significantly longer (48 hours/round) reannealing hybridization period.

Table 5

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	PE Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	PE Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	PE Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25:3389-3402.	ESTs: Probability value= 1.0E-8 or less Full Length sequences: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises at least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad. Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183:63-98; and Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs: fasta E value=1.06E-6 Assembled ESTs: fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less Full Length sequences: fastx score=100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S. and J.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	Score=1000 or greater; Ratio of Score/Strength = 0.75 or larger; and, if applicable, Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol. 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322.	Score=10-50 bits for PFAM hits, depending on individual protein families

Table 5 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, M. et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221.	Normalized quality score \geq GCG-specified "HIGH" value for that particular Prosite motif. Generally, score = 1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score = 120 or greater; Match length = 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies.	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12:431-439.	Score = 3.5 or greater
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

What is claimed is:

1. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:
 - 5 a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-5,
 - b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-5,
 - c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, and
 - 10 d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-5.
2. An isolated polypeptide of claim 1 selected from the group consisting of SEQ ID NO:1-5.
- 15 3. An isolated polynucleotide encoding a polypeptide of claim 1.
4. An isolated polynucleotide encoding a polypeptide of claim 2.
5. An isolated polynucleotide of claim 4 selected from the group consisting of SEQ ID
20 NO:6-10.
6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.
- 25 7. A cell transformed with a recombinant polynucleotide of claim 6.
8. A transgenic organism comprising a recombinant polynucleotide of claim 6.
9. A method for producing a polypeptide of claim 1, the method comprising:
 - 30 a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and
 - b) recovering the polypeptide so expressed.
- 35

10. An isolated antibody which specifically binds to a polypeptide of claim 1.

11. An isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of:

- 5 a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:6-10,
 b) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:6-10,
 c) a polynucleotide sequence complementary to a),
 d) a polynucleotide sequence complementary to b), and
10 e) an RNA equivalent of a)-d).

12. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 11.

15 13. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:

- a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization
20 complex is formed between said probe and said target polynucleotide or fragments thereof, and
 b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.

 14. A method of claim 13, wherein the probe comprises at least 60 contiguous nucleotides.
25

 15. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:

- a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and
30 b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

 16. A composition comprising an effective amount of a polypeptide of claim 1 and a pharmaceutically acceptable excipient.
35

17. A composition of claim 16, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-5.

18. A method for treating a disease or condition associated with decreased expression of functional APOP, comprising administering to a patient in need of such treatment the composition of claim 16.

19. A method for screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:

10 a) exposing a sample comprising a polypeptide of claim 1 to a compound, and

 b) detecting agonist activity in the sample.

20. A composition comprising an agonist compound identified by a method of claim 19 and a pharmaceutically acceptable excipient.

15

21. A method for treating a disease or condition associated with decreased expression of functional APOP, comprising administering to a patient in need of such treatment a composition of claim 20.

22. A method for screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:

20 a) exposing a sample comprising a polypeptide of claim 1 to a compound, and

 b) detecting antagonist activity in the sample.

23. A composition comprising an antagonist compound identified by a method of claim 22 and a pharmaceutically acceptable excipient.

25

24. A method for treating a disease or condition associated with overexpression of functional APOP, comprising administering to a patient in need of such treatment a composition of claim 23.

30

25. A method of screening for a compound that specifically binds to the polypeptide of claim 1, said method comprising the steps of:

 a) combining the polypeptide of claim 1 with at least one test compound under suitable conditions, and

35 b) detecting binding of the polypeptide of claim 1 to the test compound, thereby identifying a

compound that specifically binds to the polypeptide of claim 1.

26. A method of screening for a compound that modulates the activity of the polypeptide of claim 1, said method comprising:

- 5 a) combining the polypeptide of claim 1 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 1,
- b) assessing the activity of the polypeptide of claim 1 in the presence of the test compound, and
- c) comparing the activity of the polypeptide of claim 1 in the presence of the test compound
- 10 with the activity of the polypeptide of claim 1 in the absence of the test compound, wherein a change in the activity of the polypeptide of claim 1 in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide of claim 1.

27. A method for screening a compound for effectiveness in altering expression of a target

15 polynucleotide, wherein said target polynucleotide comprises a sequence of claim 5, the method comprising:

- a) exposing a sample comprising the target polynucleotide to a compound, and
- b) detecting altered expression of the target polynucleotide.

20 28. A method for assessing toxicity of a test compound, said method comprising:

- a) treating a biological sample containing nucleic acids with the test compound;
- b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide of claim 11 under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological
- 25 sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide of claim 11 or fragment thereof;
- c) quantifying the amount of hybridization complex; and
- d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the
- 30 amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

<110> INCYTE GENOMICS, INC.
TANG, Y. Tom
HILLMAN, Jennifer L.
YUE, Henry
BAUGHN, Mariah R.
AZIMZAI, Yalda
AU-YOUNG, Janice

<130> PF-0740 PCT

<140> To Be Assigned

<141> Herewith

<150> 60/153,403

<151> 1999-09-10

<160> 10

<170> PERL Program

<210> 1

<211> 115

<212> PRT

<213> Homo sapiens

<220>

<221> misc feature

<223> Incyte ID No: 898712CD1

<400> 1

Met	Gly	Met	Gly	Arg	Ala	Thr	Met	Ser	Pro	Gln	Gly	Leu	Gly	Ser
1				5					10					15
Ala	Glu	Met	Thr	Leu	Ala	Gln	Ser	Val	Cys	Leu	Val	Pro	Thr	Glu
				20					25					30
Leu	Lys	Pro	Pro	Met	Gly	Glu	Val	Val	Thr	Pro	Thr	Ala	Pro	Gln
				35					40					45
Glu	Pro	Pro	Glu	Ser	Val	Arg	Pro	Ser	Ala	Pro	Pro	Ala	Glu	Leu
				50					55					60
Glu	Val	Gln	Ala	Ser	Glu	Cys	Val	Val	Cys	Leu	Glu	Arg	Glu	Ala
				65					70					75
Gln	Met	Ile	Phe	Leu	Asn	Cys	Gly	His	Val	Cys	Cys	Cys	Gln	Gln
				80					85					90
Cys	Cys	Gln	Pro	Leu	Arg	Thr	Cys	Pro	Leu	Cys	Arg	Gln	Asp	Ile
				95					100					105
Ala	Gln	Arg	Leu	Arg	Ile	Tyr	His	Ser	Ser					
				110					115					

<210> 2

 $\langle 211 \rangle$ 311

<212> PRT

<213> Homo sapiens

<220>

```
<221> misc_feature
```

<223> Incyte ID No: 1301667CD1

<400> 2

Met	Gly	Leu	Lys	Glu	Gly	Tyr	Glu	Glu	Thr	Gly	Val	Leu	Thr	Leu
1				5					10					15
Val	Ile	His	Ser	Gln	Tyr	Pro	Asp	Lys	Lys	Arg	Leu	Glu	Gly	Leu
				20					25					30
Ser	Lys	Gln	Leu	Asp	Trp	Asn	Val	Arg	Lys	Ile	Gln	Cys	Trp	Phe
				35					40					45

Arg	His	Arg	Arg	Asn	Gln	Asp	Lys	Pro	Pro	Thr	Leu	Thr	Lys	Phe	
				50					55					60	
Cys	Glu	Ser	Met	Trp	Arg	Phe	Thr	Phe	Tyr	Leu	Cys	Ile	Phe	Cys	
				65					70					75	
Tyr	Gly	Ile	Arg	Phe	Leu	Trp	Ser	Ser	Pro	Trp	Phe	Trp	Asp	Ile	
				80					85					90	
Arg	Gln	Cys	Trp	His	Asn	Tyr	Pro	Phe	Gln	Pro	Leu	Ser	Ser	Gly	
				95					100					105	
Leu	Tyr	His	Tyr	Tyr	Ile	Met	Glu	Leu	Ala	Phe	Tyr	Trp	Ser	Leu	
				110					115					120	
Met	Phe	Ser	Gln	Phe	Thr	Asp	Ile	Lys	Arg	Lys	Asp	Phe	Leu	Ile	
				125					130					135	
Met	Phe	Val	His	His	Leu	Val	Thr	Ile	Gly	Leu	Ile	Ser	Phe	Ser	
				140					145					150	
Tyr	Ile	Asn	Asn	Met	Val	Arg	Val	Gly	Thr	Leu	Ile	Met	Cys	Leu	
				155					160					165	
His	Asp	Val	Ser	Asp	Phe	Leu	Leu	Glu	Ala	Ala	Lys	Leu	Ala	Asn	
				170					175					180	
Tyr	Ala	Lys	Tyr	Gln	Arg	Leu	Cys	Asp	Thr	Leu	Phe	Val	Ile	Phe	
				185					190					195	
Ser	Ala	Val	Phe	Met	Val	Thr	Arg	Leu	Gly	Ile	Tyr	Pro	Phe	Trp	
				200					205					210	
Ile	Leu	Asn	Thr	Thr	Leu	Phe	Glu	Ser	Trp	Glu	Ile	Ile	Gly	Pro	
				215					220					225	
Tyr	Ala	Ser	Trp	Trp	Leu	Leu	Asn	Gly	Leu	Leu	Leu	Thr	Leu	Gln	
				230					235					240	
Leu	Leu	His	Val	Ile	Trp	Ser	Tyr	Leu	Ile	Ala	Arg	Ile	Ala	Leu	
				245					250					255	
Lys	Ala	Leu	Ile	Arg	Gly	Lys	Val	Ser	Lys	Asp	Asp	Arg	Ser	Asp	
				260					265					270	
Val	Glu	Ser	Ser	Ser	Glu	Glu	Glu	Asp	Val	Thr	Thr	Cys	Thr	Lys	
				275					280					285	
Ser	Pro	Cys	Asp	Ser	Ser	Ser	Ser	Asn	Gly	Ala	Asn	Arg	Val	Asn	
				290					295					300	
Gly	His	Met	Gly	Gly	Ser	Tyr	Trp	Ala	Glu	Glu					
				305					310						

<210> 3
 <211> 219
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1623420CD1

<400>	3														
Met	Pro	Arg	Gly	Arg	Lys	Ser	Arg	Arg	Arg	Arg	Asn	Ala	Arg	Ala	
	1				5				10					15	
Ala	Glu	Glu	Asn	Arg	Asn	Asn	Arg	Lys	Ile	Gln	Ala	Ser	Glu	Ala	
				20					25					30	
Ser	Glu	Thr	Pro	Met	Ala	Ala	Ser	Val	Val	Ala	Ser	Thr	Pro	Glu	
				35					40					45	
Asp	Asp	Leu	Ser	Gly	Pro	Glu	Glu	Asp	Pro	Ser	Thr	Pro	Glu	Glu	
				50					55					60	
Ala	Ser	Thr	Thr	Pro	Glu	Glu	Ala	Ser	Ser	Thr	Ala	Gln	Ala	Gln	
				65					70					75	
Lys	Pro	Ser	Val	Pro	Arg	Ser	Asn	Phe	Gln	Gly	Thr	Lys	Lys	Ser	
				80					85					90	
Leu	Leu	Met	Ser	Ile	Leu	Ala	Leu	Ile	Phe	Ile	Met	Gly	Asn	Ser	
				95					100					105	
Ala	Lys	Glu	Ala	Leu	Val	Trp	Lys	Val	Leu	Gly	Lys	Leu	Gly	Met	
				110					115					120	
Gln	Pro	Gly	Arg	Gln	His	Ser	Ile	Phe	Gly	Asp	Pro	Lys	Lys	Ile	
				125					130					135	
Val	Thr	Glu	Glu	Phe	Val	Arg	Arg	Gly	Tyr	Leu	Ile	Tyr	Lys	Pro	
				140					145					150	

Val	Pro	Arg	Ser	Ser	Pro	Val	Glu	Tyr	Glu	Phe	Phe	Trp	Gly	Pro
				155					160					165
Arg	Ala	His	Val	Glu	Ser	Ser	Lys	Leu	Lys	Val	Met	His	Phe	Val
				170					175					180
Ala	Arg	Val	Arg	Asn	Arg	Cys	Ser	Lys	Asp	Trp	Pro	Cys	Asn	Tyr
				185					190					195
Asp	Trp	Asp	Ser	Asp	Asp	Asp	Ala	Glu	Val	Glu	Ala	Ile	Leu	Asn
				200					205					210
Ser	Gly	Ala	Arg	Gly	Tyr	Ser	Ala	Pro						
				215										

<210> 4
 <211> 358
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2240612CD1

<400> 4

Met	Ala	Ala	Val	Leu	Lys	Pro	Val	Leu	Leu	Gly	Leu	Arg	Asp	Ala
1				5					10					15
Pro	Val	His	Gly	Ser	Pro	Thr	Gly	Pro	Gly	Ala	Trp	Thr	Ala	Ser
				20					25					30
Lys	Leu	Gly	Gly	Ile	Pro	Asp	Ala	Leu	Pro	Thr	Val	Ala	Ala	Pro
				35					40					45
Arg	Pro	Val	Cys	Gln	Arg	Cys	Gly	Gln	Pro	Leu	Ala	Leu	Val	Val
				50					55					60
Gln	Val	Tyr	Cys	Pro	Leu	Glu	Gly	Ser	Pro	Phe	His	Arg	Leu	Leu
				65					70					75
His	Val	Phe	Ala	Cys	Ala	Cys	Pro	Gly	Cys	Ser	Thr	Gly	Gly	Ala
				80					85					90
Arg	Ser	Trp	Lys	Val	Phe	Arg	Ser	Gln	Cys	Leu	Gln	Val	Pro	Glu
				95					100					105
Arg	Glu	Ala	Gln	Asp	Ala	Gln	Lys	Gln	Gly	Asn	Ser	Leu	Ala	Ala
				110					115					120
Glu	Asp	Trp	Cys	Glu	Gly	Ala	Asp	Asp	Trp	Gly	Ser	Asp	Thr	Glu
				125					130					135
Glu	Gly	Pro	Ser	Pro	Gln	Phe	Thr	Leu	Asp	Phe	Gly	Asn	Asp	Ala
				140					145					150
Ser	Ser	Ala	Lys	Asp	Val	Asp	Trp	Thr	Ala	Arg	Leu	Gln	Asp	Leu
				155					160					165
Arg	Leu	Gln	Asp	Ala	Val	Leu	Gly	Ala	Ala	His	Pro	Val	Pro	Pro
				170					175					180
Gly	Leu	Pro	Leu	Phe	Leu	Pro	Tyr	Tyr	Ile	Cys	Val	Ala	Asp	Glu
				185					190					195
Asp	Asp	Tyr	Arg	Asp	Phe	Val	Asn	Leu	Asp	His	Ala	His	Ser	Leu
				200					205					210
Leu	Arg	Asp	Tyr	Gln	Gln	Arg	Glu	Gly	Ile	Ala	Met	Asp	Gln	Leu
				215					220					225
Leu	Ser	Gln	Ser	Leu	Pro	Asn	Asp	Gly	Asp	Glu	Lys	Tyr	Glu	Lys
				230					235					240
Thr	Ile	Ile	Lys	Ser	Gly	Asp	Gln	Thr	Phe	Tyr	Lys	Phe	Met	Lys
				245					250					255
Arg	Ile	Ala	Ala	Cys	Gln	Glu	Gln	Ile	Leu	Arg	Tyr	Ser	Trp	Ser
				260					265					270
Gly	Glu	Pro	Leu	Phe	Leu	Thr	Cys	Pro	Thr	Ser	Glu	Val	Thr	Glu
				275					280					285
Leu	Pro	Ala	Cys	Ser	Gln	Cys	Gly	Gly	Gln	Arg	Ile	Phe	Glu	Phe
				290					295					300
Gln	Leu	Met	Pro	Ala	Leu	Val	Ser	Met	Leu	Lys	Ser	Ala	Asn	Leu
				305					310					315
Gly	Leu	Ser	Val	Glu	Phe	Gly	Thr	Ile	Leu	Val	Tyr	Thr	Cys	Glu
				320					325					330
Lys	Ser	Cys	Trp	Pro	Pro	Asn	His	Gln	Thr	Pro	Met	Glu	Glu	Phe
				335					340					345

Cys Ile Ile Gln Glu Asp Pro Asp Glu Leu Leu Phe Lys
 350 355

<210> 5
 <211> 269
 <212> PRT
 <213> Homo sapiens
 <220>
 <221> misc_feature
 <223> Incyte ID No: 4270927CD1

<400> 5
 Met Ala Thr Glu Ser Thr Pro Ser Glu Ile Ile Glu Arg Glu Arg
 1 5 10 15
 Lys Lys Leu Leu Glu Ile Leu Gln His Asp Pro Asp Ser Ile Leu
 20 25 30
 Asp Thr Leu Thr Ser Arg Arg Leu Ile Ser Glu Glu Glu Tyr Glu
 35 40 45
 Thr Leu Glu Asn Val Thr Asp Leu Leu Lys Lys Ser Arg Lys Leu
 50 55 60
 Leu Ile Leu Val Gln Lys Lys Gly Glu Ala Thr Cys Gln His Phe
 65 70 75
 Leu Lys Cys Leu Phe Ser Thr Phe Pro Gln Ser Ala Ala Ile Cys
 80 85 90
 Gly Leu Arg His Glu Val Leu Lys His Glu Asn Thr Val Pro Pro
 95 100 105
 Gln Ser Met Gly Ala Ser Ser Asn Ser Glu Asp Ala Phe Ser Pro
 110 115 120
 Gly Ile Lys Gln Pro Glu Ala Pro Glu Ile Thr Val Phe Phe Ser
 125 130 135
 Glu Lys Glu His Leu Asp Leu Glu Thr Ser Glu Phe Phe Arg Asp
 140 145 150
 Lys Lys Thr Ser Tyr Arg Glu Thr Ala Leu Ser Ala Arg Lys Asn
 155 160 165
 Glu Lys Glu Tyr Asp Thr Pro Glu Val Thr Leu Ser Tyr Ser Val
 170 175 180
 Glu Lys Val Gly Cys Glu Val Pro Ala Thr Ile Thr Tyr Ile Lys
 185 190 195
 Asp Gly Gln Arg Tyr Glu Glu Leu Asp Ser Leu Tyr Leu Gly
 200 205 210
 Lys Glu Glu Tyr Leu Gly Ser Val Asp Thr Pro Glu Asp Ala Glu
 215 220 225
 Ala Thr Val Glu Glu Glu Val Tyr Asp Asp Pro Glu His Val Gly
 230 235 240
 Tyr Asp Gly Glu Glu Asp Phe Glu Asn Ser Glu Thr Thr Glu Phe
 245 250 255
 Ser Gly Glu Glu Pro Ser Tyr Glu Gly Ser Glu Thr Ser Leu
 260 265

<210> 6
 <211> 1153
 <212> DNA
 <213> Homo sapiens
 <220>
 <221> misc_feature
 <223> Incyte ID No: 898712CB1

<400> 6
 aagcgctcccggggacattag agatcagagg acccacttaa caagcgctgc ggccatccct 60
 gcagagagag acttcactcc cattctgcag atgggaaatg aggctcagag aagagaagaa 120
 gagagtgtgt ggcaaggaga gcacttccct cagatctgag agcagaggct ccccccacct 180
 ccagctcact caccatgggg atgggcaggg ccacaatgag cccccagggg ttagggtcag 240
 cggagatgac cctggctcag tctgtctgtc tgggtccccac agagctgaaa ccaccaatgg 300
 gtgaggtcgt caccctacg gccccccagg agcctcctga gtctgtgagg ccatccgctc 360
 cccctgcaga gctggagggtg caggcctcag agtgtgtcgt gtgcctggaa cgggaggccc 420

```

agatgatctt cctcaactgt ggccacgtct gctgctgcca gcagtgtctgc cagccactgc 480
gcacctgccc gctgtgccc caggacatcg cccagcgcc cgcacatctac cacagcagct 540
gagtgtctgcc cgcccacctg ggccctgtcc tagccctgcc tcggccactg tgagccccgg 600
gctcctgtct agccttgtgc cagccagact cgtatgaggg tccccctgc cctgggcccc 660
ttccccactg cccaggagcc cccatcctaa gctccaagca tgtctggggc aggcagaggt 720
gctcctcatc catgacacca ccagtctgaa tggctcctggg ggctggggct ggagaggccg 780
ctgcaccacc acccgagcct gggagccagc gtcccagcct aatcacggat ctgctgcctc 840
ccagctgtct tgactgaagg ccaccgcccc tgcaggagct tgggtcctca tctgggggcc 900
atgcacaggc ccgtcccacc ctgcatgtgg gaagggagca ggagggcctg gctgggtgag 960
gggaggcctt cctgggaagg cgtgtgtgtc aggcctgtgc tcacagtggc accagcaacc 1020
ctgggtctcc ctctctgtct cccccagaa ccccgggggc ctctgtctct ccacaactgt 1080
ccctccttac cccatgtagc tcgatccgaa gcaggagtgt caataaacct gtcttcagt 1140
cgaaaaaaa aaa

```

<210> 7
 <211> 2036
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1301667CB1

```

<400> 7
cagtactgat agcctttgca gtttgggttt tttgtttttt ggagagacaa ggttttgcca 60
tgttgcccag gctgggtctcg gactcctggg ctcaagccat ccttcggccc tcagccttca 120
aaaatcccag gattacaggc atgagccacc atgcctggcc tgcaatttgg gtttttaata 180
tcctttttta ctacagcttgg gctggaggaa gagacaaaac aaggatatgt ataggtgagc 240
agtacaggtc cttgatattg aagataaaat ttcttactaa aaatgacttt gggataaaac 300
ttggaaaagat aaacatcact tgagccactt atggctatca taccttgtcg ctacagtagt 360
caatgggatt gaaggagggg tatgaggaga cgggtgttct gactctgggt atacattcac 420
agtatcctga taagaaaagg ctggagggcc tgtcaaagca gctggattgg aatgtccgaa 480
aaatccaatg ctggttttcg catcggagga atcaggacaa gcccccaacg cttactaaat 540
tctgtgaaag catgtggaga ttacattttt atttatgtat attctgctat ggaattagat 600
ttctctggtc gtcaccttgg ttctgggaca tccgacagt ctggcacaa taaccatttc 660
agcctctttc aagtgggctt tatcactatt atatcatgga attggccttc tatttggtccc 720
ttatgttttc tcagttttaca gacattaaaa gaaaggactt cctgatcatg tttgtgcatc 780
acttggtcac cattgggctt atctccttct cctacatcaa caatatgggt cgagtgggaa 840
ctctgatcat gtgtctacat gatgtctcag atttcttgcg ggaggcagcc aaactggcca 900
attatgccaa gtatcagcgg ctctgtgaca ccttttttgt gatcttcagt gctgttttta 960
tggttacacg actaggaatc tatccattct ggattctgaa cacgaccctc tttgagagtt 1020
gggagataat cgggccttat gcttcatggt ggctcctcaa tggcctgctg ctgaccctac 1080
agcttctgca tgtcatctgg tcctacctaa ttgcacggat agtgaggag cagctcagag gaagaagatg 1200
ggggaaagg atcgaaggat gatcgagtg atgtggagag cagctcagag aatcggtgcc aatcggtgga 1260
tgaccacctg cacaaaaagt cctgtgaca gtagctccag caatggtgcc agggacttca 1320
atggtcacat gggaggcagc tactgggctg aagagtaagg tggttgctat agggacttca 1380
gcacacatgg acttgtaggg ccactggcaa catactctc ttggcccttc ccatatctac 1440
tcttctgtga ttgggagact gcaaggcact gaggagtatc aaagaagcaa atattttcac 1500
tttgaaagaa aactgccatt ttgtatttaa tagccttcag gttctttcag taatgttatt 1560
tgctctgtgt gtttttgtgt gtttgttgat gtgcgtttgt gcatatgcgt gagtttcatt 1620
gccgggggtg gggcacaatt gtggactggg gccatgaggg cttccctggt cccactgaa 1680
cccaccttag ttccacattt ggctgcactt tgaattatgc cgactccaga cttctcctcc 1740
ttttttgccc ttggctcttg acactctaaa cccctggacc atctgaatgg agcagccaag 1800
ttcagtccca catttctgta ctgttctctt ttcacagctg gaatatgtca catgatgaag 1860
ttgtatagaa acagaacat ggatggatgg ccaggattgc cgtgggtcct agctagatcc 1920
ccttcctatc aatcacctga tagcaacagg gacagctgcc aataccctgc tctttactca 1980
atgggtacca gggaggagc atgggaagag ggtgagctga gggctggagg agggcaacag 2036
ccactgggtg agctgttcac ggtcttatac tattgtttgt gattaaaagt gcttca

```

<210> 8
 <211> 1433
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1623420CB1

<400> 8
cgcgggcttg cagctgcggc aagtgcctggc ggcggtctgt cgcgcaagtc agctggcggtg 60
ggaactaccc tttgtagctg agaacggctt gtttattgct acaaagactc tattgacatt 120
ggtagcttca gcggcagcag cttcttacgg tataaagctg ttgcttctctg aagaggctac 180
aagcatcctt ccctaggact gctgtaagct ttgagcctct agcaggagac atgcctcggg 240
gacgaaagag tcggcgccgc cgtaatgcga gagccgcaga agagaaccgc aacaatcgca 300
aaatccaggc ctgagaggcc tccgagacc ctatggccgc ctctgtggta gcgagcacc 360
ccgaagacga cctgagcggc ccggaggaag acccgagcac tccagaggag gcctctacca 420
cccctgaaga agcctcgagc actgcccga cacaaaagcc ttcagtggcc cggagcaatt 480
ttcagggcac caagaaaagt ctctgatgt ctatattagc gctcatcttc atcatgggca 540
acagcgccaa ggaagctctg gtctggaaag tgctggggaa gttaggaatg cagcctggac 600
gtcagcacag catctttgga gatccgaaga agatcgtcac agaagagttt gtgcgagag 660
ggtagctgat ttataaacgg gtgccccgta ccaggtccgt ggagtatgag ttcttctggg 720
ggccccgagc acacgtggaa tcgagcaaac tgaaagtcac gcattttgtg gcaagggttc 780
gtaaccgatg ctctaaagac tggccttgta attatgactg ggattcggac gatgatgcag 840
agggtgaggc tatcctcaat tcaggtgcta ggggttattc cgccccctaa gtagatctga 900
ggcagaccct tgggggtgta aaagagagtc agcaggtacc caaggagtat atgccagggt 960
cctaagttga aaatgatgtc gattgggggc gggggacact gtatttgata tttgtgatca 1020
gtgatcattg ttcaactgcg aaatagagtg tttgcttttg ataattgaaa attgtattcg 1080
ttttaaaatt ccgtttgttg agaataacaa tatgtttaaa aatataattg aacaaatttt 1140
tttctttgtt tcctgtcatt gacatttagt ataacagttt tgctaacgtt ctaaaatgaa 1200
gtcgttccat cataatctat gatcttgtac agcacttata gaaataagct gttcttttga 1260
agttgaaata cccagtaaaa tgttgaagaa ggatggagga tttcttcata tctgacgttt 1320
ctgaaaccct ttgtgtctgc tgttgtgtga agattgacat ttaccatgat tttccttagt 1380
tactgcagaa catagagaaa aataaaagcc taacgaatag taaaaaaaaa aaa 1433

<210> 9
<211> 1166
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2240612CB1

<400> 9
cccacgcgtc cgctggctgc ccggcggcca tggcgccgct tctgaagccg gtgctgctgg 60
gccttcgaga tgcgcgggtg cagggcagcc ccacagggcc ggggtgctgg actgctagca 120
agctggggcg cattccggat gctctgccca ccgtggctgc gccaggcccc gtgtgtcagc 180
gctgcgggca gccgctcgct ctggctcgtg aggtgtattg cccgctggaa ggctccccgt 240
ttcacctgtc gctgcacgtg ttcgctgctg cctgccccgg ctgtagcacc ggcgggtgcgc 300
gtagctggaa ggtgttcgcg tcccagtgcc tgcaggtgcc agagagagag gcgcaggacg 360
ctcagaaaca gggaaacagc cttgcagctg aggactgggt tgaaggtgct gatgactggg 420
gaagtgatac tgaggagggg ccttcaccac agtttacctt ggattttggg aatgatgcca 480
gcagtgccaa agacgtagac tggactgtc ggctccaaga cctccgctg caggatgctg 540
tcctgggtgc tgccatcct gtgctcctg ggctgcccgt ctctctgcc tactacatct 600
gtgttgcaga tgaggatgat tacagggaact ttgtcaacct ggatcatgcc cacagccttc 660
tgagggacta tcagcagaga gaaggcattg ccattggatca gttgctttcc caaagccttc 720
ctaattgatg tgatgaaaaa tatgagaaga ccataattaa aagtggagat cagacgtttt 780
acaaattcat gaagcgaatt gctgcttctc aggagcagat tttgaggat tcttgaggatg 840
gagagccact ctttttgacc tgccctacat cagaagtcac cgagctccca gcctgcagcc 900
agtgtggagg ccaaaggata tttgagtttc agcttatgcc agcactgggt agcatgctca 960
agagtgtctaa tttaggtctt tctgtggaat ttggaacaat tctagttac acatgtgaga 1020
agagttgctg gcccccaat catcagactc ccattggaaga attttgtatt atacaagaag 1080
accagatga attattgttt aagtagagca tttcctttta ttaatatata ttaaaacaaa 1140
tgtttacatc caaaaaaaaa aaaaaa 1166

<210> 10
<211> 824
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4270927CB1

<400> 10

gaacaatggc	taccgagagt	actccctcag	agatcataga	aagagaaaga	aaaaagttgc	60
ttgaaatcct	tcaacatgat	cctgattcta	tcttagacac	gttaacttct	cgagggtga	120
tttctgagga	agagtatgag	actctggaga	atgttacaga	tctcctgaag	aaaagtcgga	180
agctgttaat	tttggtagag	aaaaagggag	aggcgacctg	tcagcatttt	ctcaagtgtt	240
tatttagtac	ttttccacag	tcagctgccg	tttgcggctt	aaggcatgaa	gttttaaaac	300
atgagaatac	agtacctcct	caatctatgg	gggcaagcag	taattcagaa	gatgcttttt	360
ctcctggaat	aaaacagcct	gaagcccctg	agatcacagt	gttcttcagt	gagaaggaac	420
acttggaattt	ggaaacctct	gagtttttca	gggacaagaa	aactagtatt	agggaaacag	480
ctttgtctgc	caggaagaat	gagaaggaat	atgacacacc	agaagtcaca	ttatcatatt	540
cagttgagaa	agttggatgt	gaagttccag	caactattac	atatataaaa	gatggacaga	600
gatatgagga	gctagatgat	tctttatact	taggaaaaga	ggaatatcta	ggatctgttg	660
acaccctga	agatgcagaa	gccactgtgg	aagaggaggt	ttatgatgac	ccagagcacg	720
ttggatatga	tggatgaagag	gacttcgaga	attcagaaac	cacagagttc	tctggtgaag	780
aaccaagtta	tgagggatca	gaaaccagcc	tttaattgga	ggag		824

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
15 March 2001 (15.03.2001)

PCT

(10) International Publication Number
WO 01/18042 A3

(51) International Patent Classification⁷: **C12N 15/12**,
C07K 14/47, 16/18, C12N 5/10, C12Q 1/68, A61K 38/17,
G01N 33/68, A01K 67/027

Road, San Leandro, CA 94577 (US). **AZIMZAI**, Yalda
[US/US]; 2045 Rock Springs, Hayward, CA 94545 (US).
AU-YOUNG, Janice [US/US]; 233 Golden Eagle Lane,
Brisbane, CA 94005 (US).

(21) International Application Number: PCT/US00/24483

(22) International Filing Date:
6 September 2000 (06.09.2000)

(74) Agents: **HAMLET-COX**, Diana et al.; Incyte Genomics,
Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).

(25) Filing Language: English

(26) Publication Language: English

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ,
DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR,
HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR,
LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ,
NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM,
TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(30) Priority Data:
60/153,403 10 September 1999 (10.09.1999) US

(84) Designated States (*regional*): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG,
CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(63) Related by continuation (CON) or continuation-in-part
(CIP) to earlier application:
US 60/153,403 (CIP)
Filed on 10 September 1999 (10.09.1999)

(71) Applicant (*for all designated States except US*): **INCYTE
GENOMICS, INC.** [US/US]; 3160 Porter Drive, Palo
Alto, CA 94304 (US).

Published:
— with international search report

(72) Inventors: and

(75) Inventors/Applicants (*for US only*): **TANG**, Tom, Y.
[CN/US]; 4230 Ranwick Court, San Jose, CA 95118
(US). **HILLMAN**, Jennifer, L. [US/US]; 230 Monrow
Drive, #17, Mountain View, CA 94040 (US). **YUE**,
Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087
(US). **BAUGHN**, Mariah, R. [US/US]; 14244 Santiago

(88) Date of publication of the international search report:
6 December 2001

*For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.*

WO 01/18042 A3

(54) Title: APOPTOSIS PROTEINS

(57) Abstract: The invention provides human apoptosis proteins (APOP) and polynucleotides which identify and encode APOP. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with expression of APOP.

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 00/24483

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/12 C07K14/47 C07K16/18 C12N5/10 C12Q1/68
A61K38/17 G01N33/68 A01K67/027

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C07K C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, EMBL, WPI Data, BIOSIS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	<p>DATABASE EMBL [Online] 24 June 1997 (1997-06-24) HILLIER L. ET AL.: "zt34e05.s1 Soares ovary tumor NbHOT Homo sapiens cDNA clone IMAGE:724256" retrieved from EBI Database accession no. AA482565 XP002160354 abstract</p> <p style="text-align: center;">---</p> <p style="text-align: center;">-/--</p>	1-27

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "8" document member of the same patent family

Date of the actual completion of the international search

15 February 2001

Date of mailing of the international search report

17.05.01

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Gurdjian, D

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 00/24483

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	DATABASE EMBL [Online] 21 November 1997 (1997-11-21) HILLIER L. ET AL.: "af79d05.r1 Soares NhHMPu S1 Homo sapiens cDNA clone 1048233 5'." retrieved from EBI Database accession no. AA670104 XP002160355 abstract	1-27
A	--- KOSEKI TAKEYOSHI ET AL: "ARC, an inhibitor of apoptosis expressed in skeletal muscle and heart that interacts selectively with caspases" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA,US,NATIONAL ACADEMY OF SCIENCE. WASHINGTON, vol. 95, no. 9, 28 April 1998 (1998-04-28), pages 5156-5160, XP002137398 ISSN: 0027-8424 cited in the application the whole document	1-27
P,X	--- WO 00 29430 A (ZYMOGENETICS INC) 25 May 2000 (2000-05-25) page 94 -page 96; claims 1-21 -----	1,3,7, 9-16, 25-27

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 00/24483

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Although claims 18,21,24 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.
2. ☒ Claims Nos.: 20 23
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
see FURTHER INFORMATION sheet PCT/ISA/210
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-28 partially

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-28 partially

Polypeptides with amino acid sequence with seq.id.1 and corresponding nucleotide sequence with seq.id.6, transformed cells , transgenic organism , method of producing a protein , isolated antibody , method for detecting a target , pharmaceutical composition , method of screening a compound ,and method of assessing toxicity .

2. Claims: 1-28 partially

Polypeptides with amino acid sequence with seq.id.2 and corresponding nucleotide sequence with seq.id.7, transformed cells , transgenic organism , method of producing a protein , isolated antibody , method for detecting a target , pharmaceutical composition , method of screening a compound ,and method of assessing toxicity .

3. Claims: 1-28 partially

Polypeptides with amino acid sequence with seq.id.3 and corresponding nucleotide sequence with seq.id.8, transformed cells , transgenic organism , method of producing a protein , isolated antibody , method for detecting a target , pharmaceutical composition , method of screening a compound ,and method of assessing toxicity .

4. Claims: 1-28 partially

Polypeptides with amino acid sequence with seq.id.4 and corresponding nucleotide sequence with seq.id.9, transformed cells , transgenic organism , method of producing a protein , isolated antibody , method for detecting a target , pharmaceutical composition , method of screening a compound ,and method of assessing toxicity .

5. Claims: 1-28 partially

Polypeptides with amino acid sequence with seq.id.5 and corresponding nucleotide sequence with seq.id.10, transformed cells , transgenic organism , method of producing a protein , isolated antibody , method for detecting a target , pharmaceutical composition , method of screening a compound ,and method

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

of assessing toxicity .

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 20 23

Claim 20 relating to agonists and claim 23 relating to antagonists to the polypeptide of claim 1 could not be searched as its subject-matter was insufficiently disclosed .

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

Information on patent family members

PCT/US 00/24483

Form PCT/ISA/210 (patent family annex) (July 1992)